Table 2.13.3.12

Accuracy and Consistency of Classification Indices: Writ (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.789	0.706	0.524
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.902	0.826
	2	0.505	0.379
	3	0.863	0.816
	4	0.663	0.537
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.960	0.013	0.028	0.944
	2/3	0.935	0.035	0.029	0.905
	3/4	0.893	0.049	0.058	0.850
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

2.13.4 Speaking

Table 2.13.4.0

Accuracy and Consistency of Classification Indices: Spek (Grade K) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.448	0.463	0.348
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.687	0.625
	2	0.631	0.493
	3	0.351	0.256
	4	0.468	0.324
	5	0.242	0.238
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.885	0.085	0.030	0.849
	2/3	0.927	0.029	0.044	0.898
	3/4	0.947	0.027	0.026	0.922
	4/5	0.950	0.025	0.025	0.923
	5/6	0.711	0.289	0.000	0.786

Table 2.13.4.1

Accuracy and Consistency of Classification Indices: Spek (Grade 1) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.671	0.564	0.438
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.801	0.686
	2	0.721	0.614
	3	0.651	0.542
	4	0.616	0.526
	5	0.522	0.398
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.947	0.026	0.027	0.922
	2/3	0.899	0.043	0.058	0.860
	3/4	0.903	0.042	0.055	0.862
	4/5	0.930	0.055	0.015	0.908
	5/6	0.985	0.015	0.000	0.985

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.2

Accuracy and Consistency of Classification Indices: Spek (Grade 2) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.666	0.565		0.432	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.829		0.725	
	2	0.641		0.519	
	3	0.783		0.693	
	4	0.527		0.448	
	5	0.394		0.340	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.958	0.021	0.021	0.938
	2/3	0.903	0.055	0.042	0.864
	3/4	0.895	0.029	0.076	0.854
	4/5	0.923	0.069	0.009	0.909
	5/6	0.969	0.031	0.000	0.969

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.3

Accuracy and Consistency of Classification Indices: Spek (Grade 3) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.663	0.558		0.420	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.807		0.698	
	2	0.644		0.519	
	3	0.762		0.669	
	4	0.534		0.478	
	5	-		-	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.956	0.024	0.020	0.933
	2/3	0.901	0.051	0.048	0.862
	3/4	0.890	0.029	0.081	0.848
	4/5	0.916	0.084	0.000	0.911
	5/6	0.955	0.045	0.000	0.955

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.4

Accuracy and Consistency of Classification Indices: Spek (Grade 4) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.618	0.516	0.380
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.831	0.725
	2	0.576	0.445
	3	0.666	0.551
	4	0.676	0.558
	5	0.419	0.363
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.966	0.016	0.017	0.950
	2/3	0.932	0.034	0.033	0.905
	3/4	0.902	0.039	0.059	0.865
	4/5	0.879	0.053	0.067	0.825
	5/6	0.927	0.073	0.000	0.923

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.5

Accuracy and Consistency of Classification Indices: Spek (Grade 5) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.623	0.523	0.382
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.843	0.740
	2	0.515	0.390
	3	0.681	0.567
	4	0.651	0.548
	5	0.448	0.381
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.963	0.017	0.020	0.945
	2/3	0.930	0.037	0.032	0.902
	3/4	0.899	0.037	0.065	0.861
	4/5	0.874	0.066	0.060	0.824
	5/6	0.944	0.056	0.000	0.941

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.6

Accuracy and Consistency of Classification Indices: Spek (Grade 6) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.624	0.521	0.398	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.808	0.701	
	2	0.589	0.465	
	3	0.640	0.531	
	4	0.677	0.559	
	5	0.423	0.363	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.953	0.025	0.022	0.931
	2/3	0.917	0.037	0.046	0.886
	3/4	0.903	0.038	0.059	0.864
	4/5	0.907	0.042	0.052	0.862
	5/6	0.934	0.066	0.000	0.928

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.7

Accuracy and Consistency of Classification Indices: Spek (Grade 7) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.617	0.520	0.391	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.825	0.730	
	2	0.553	0.429	
	3	0.609	0.497	
	4	0.734	0.618	
	5	0.298	0.250	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.951	0.025	0.024	0.929
	2/3	0.923	0.034	0.043	0.894
	3/4	0.902	0.047	0.052	0.863
	4/5	0.907	0.037	0.056	0.859
	5/6	0.919	0.081	0.000	0.916

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.8

Accuracy and Consistency of Classification Indices: Spek (Grade 8) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.636	0.536		0.419	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.865		0.783	
	2	0.460		0.349	
	3	0.648		0.536	
	4	0.640		0.527	
	5	0.497		0.426	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.943	0.025	0.033	0.918
	2/3	0.921	0.038	0.040	0.889
	3/4	0.907	0.041	0.052	0.869
	4/5	0.907	0.055	0.038	0.866
	5/6	0.944	0.056	0.000	0.935

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.9

Accuracy and Consistency of Classification Indices: Spek (Grade 9) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.648	0.559		0.445	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.878		0.812	
	2	0.477		0.367	
	3	0.693		0.587	
	4	0.576		0.456	
	5	0.396		0.349	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.933	0.033	0.034	0.904
	2/3	0.913	0.044	0.043	0.880
	3/4	0.917	0.032	0.051	0.883
	4/5	0.930	0.046	0.024	0.899
	5/6	0.935	0.065	0.000	0.932

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.10

Accuracy and Consistency of Classification Indices: Spek (Grade 10) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.665	0.585	0.475
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.887	0.822
	2	0.465	0.354
	3	0.689	0.586
	4	0.738	0.620
	5	0.229	0.200
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.936	0.028	0.036	0.908
	2/3	0.920	0.041	0.038	0.888
	3/4	0.914	0.045	0.041	0.880
	4/5	0.958	0.020	0.022	0.930
	5/6	0.925	0.075	0.000	0.935

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.11

Accuracy and Consistency of Classification Indices: Spek (Grade 11) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.662	0.582	0.469
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.879	0.805
	2	0.453	0.343
	3	0.762	0.673
	4	0.664	0.534
	5	0.234	0.215
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.939	0.027	0.034	0.911
	2/3	0.915	0.047	0.038	0.880
	3/4	0.917	0.036	0.047	0.882
	4/5	0.960	0.024	0.016	0.936
	5/6	0.918	0.082	0.000	0.929

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 2.13.4.12

Accuracy and Consistency of Classification Indices: Spek (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.672	0.597	0.485
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.850	0.772
	2	0.547	0.432
	3	0.779	0.698
	4	0.648	0.518
	5	0.227	0.224
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.931	0.033	0.036	0.902
	2/3	0.910	0.043	0.046	0.875
	3/4	0.934	0.027	0.039	0.904
	4/5	0.971	0.020	0.009	0.954
	5/6	0.917	0.083	0.000	0.930

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3 Analyses of Composite Scores

Four composite scores are calculated for ACCESS Online: Oral Language, Literacy, Comprehension, and Overall. Composite scores are calculated as weighted averages of domain scale scores, as follows:

- Oral Language: 50% Listening + 50% Speaking
- Literacy: 50% Reading + 50% Writing
- Comprehension: 30% Listening + 70% Reading
- Overall Composite: 15% Listening + 15% Speaking + 35% Reading + 35% Writing

This weighting resulted from a policy decision by the WIDA Board before the first operational administration of ACCESS, based on the view that literacy skills are paramount in developing academic language proficiency.

3.1 Scale Score Distribution for Composites

Figures and tables in this section provide scale score distributions for each of the composites, for each grade-level cluster.

For each cluster, the figure shows the distribution of the scale scores for the composite. Scale scores are plotted on the horizontal axis, grouped into units of five scale score points (e.g., 100-104, 105-109, 110-114, etc.). The number of students with scale scores falling into each range is plotted on the vertical axis.

Each table shows, by grade and by total for the grade-level cluster:

- The number of students in the analyses (count)
- The minimum observed scale score
- The maximum observed scale score
- The mean (average) scale score
- The standard deviation (std. dev.) of the scale scores

In the tables and figures in this section, scale scores which were computed using mode-adjusted scoring tables are excluded from the analysis.

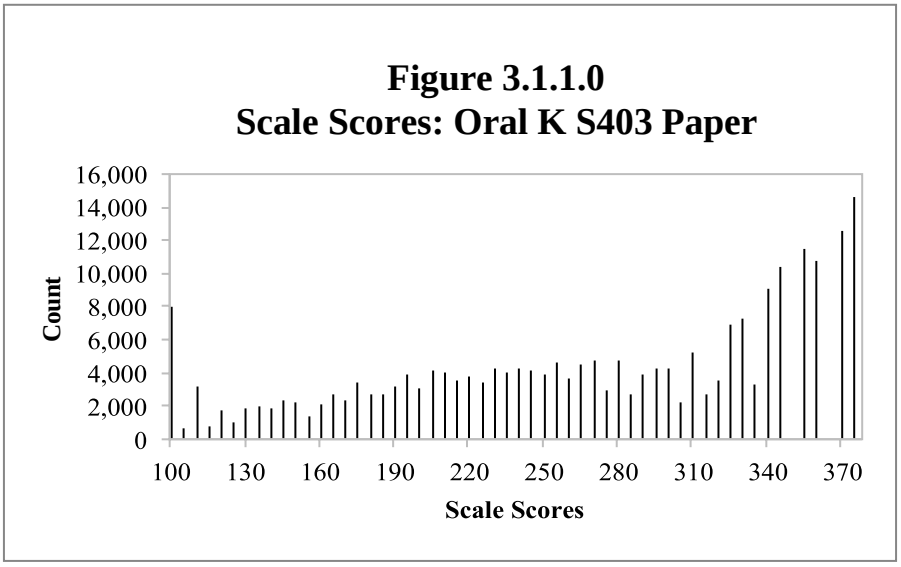
3.1.1 Oral

3.1.1.0 Kindergarten

Table 3.1.1.0

Scale Score Descriptive Statistics: Oral K S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	232,462	100	378	272.03	81.50



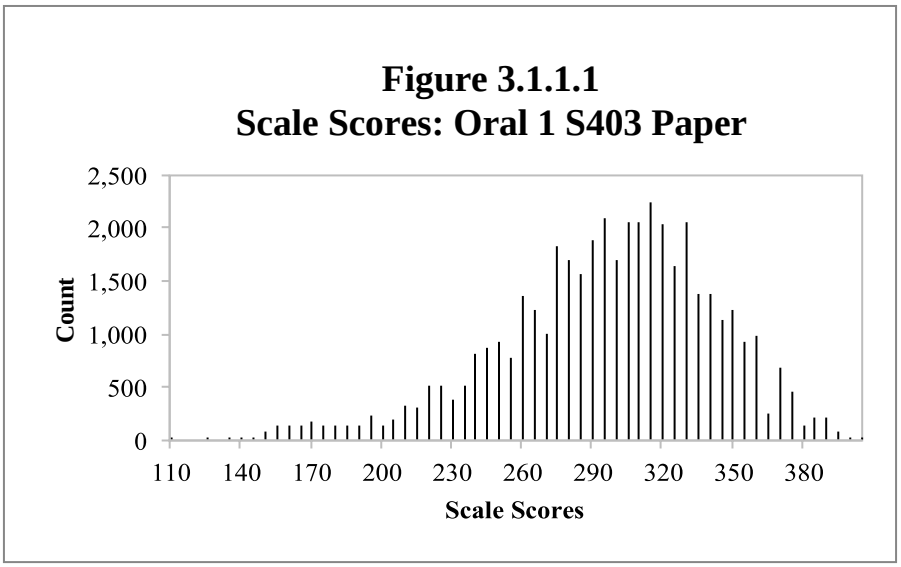
3.1.1.1 Grade 1

Table 3.1.1.1

Scale Score Descriptive Statistics: Oral 1 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	43,343	114	406	298.47	45.58

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

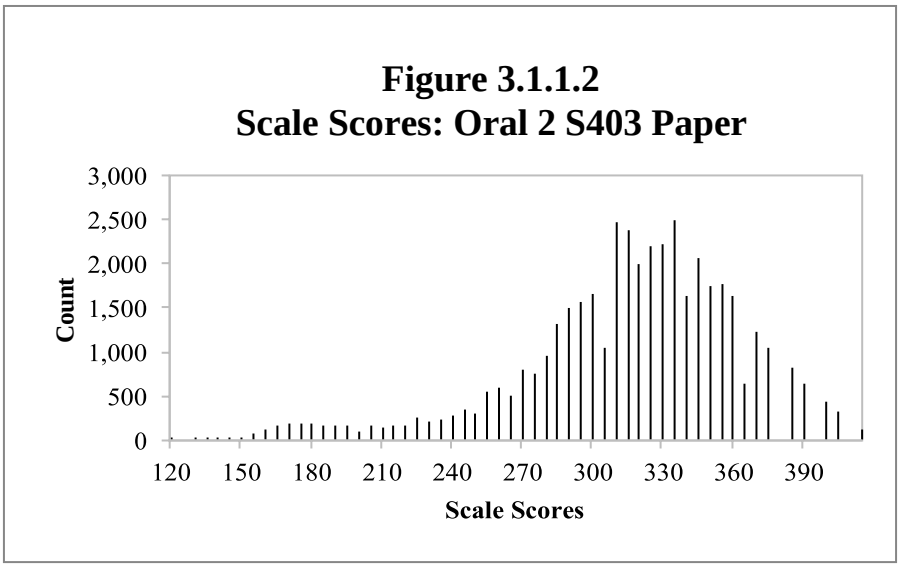
3.1.1.2 Grade 2

Table 3.1.1.2

Scale Score Descriptive Statistics: Oral 2 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	42,856	120	415	317.25	47.04

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

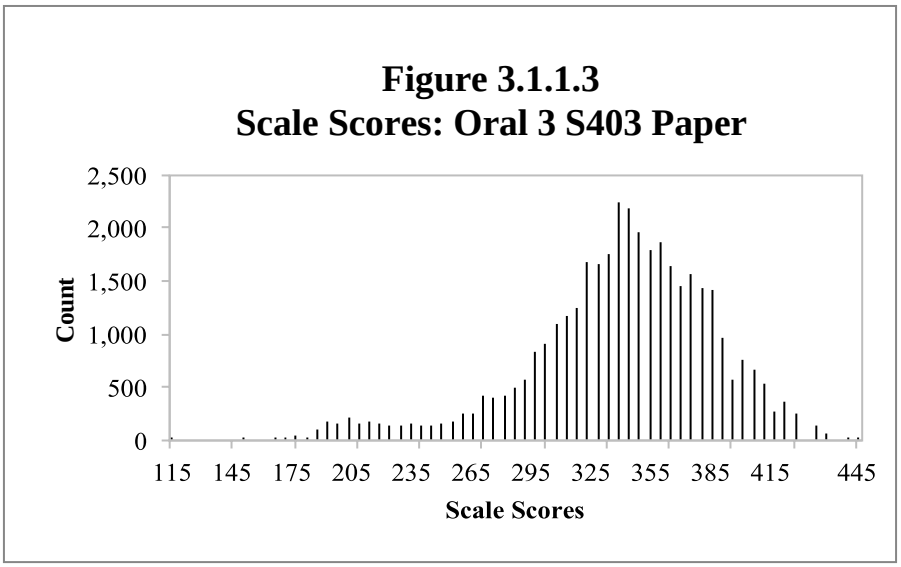
3.1.1.3 Grade 3

Table 3.1.1.3

Scale Score Descriptive Statistics: Oral 3 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	37,594	115	448	332.79	44.94

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

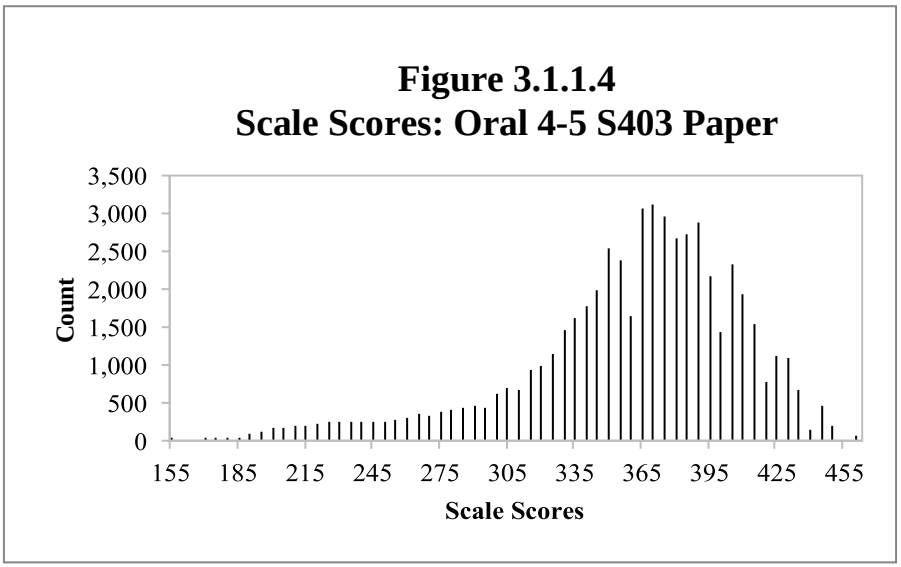
3.1.1.4 Grades 4–5

Table 3.1.1.4

Scale Score Descriptive Statistics: Oral 4-5 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	29,607	157	461	357.34	49.06
5	26,194	157	461	365.91	50.33
Total	55,801	157	461	361.36	49.84

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

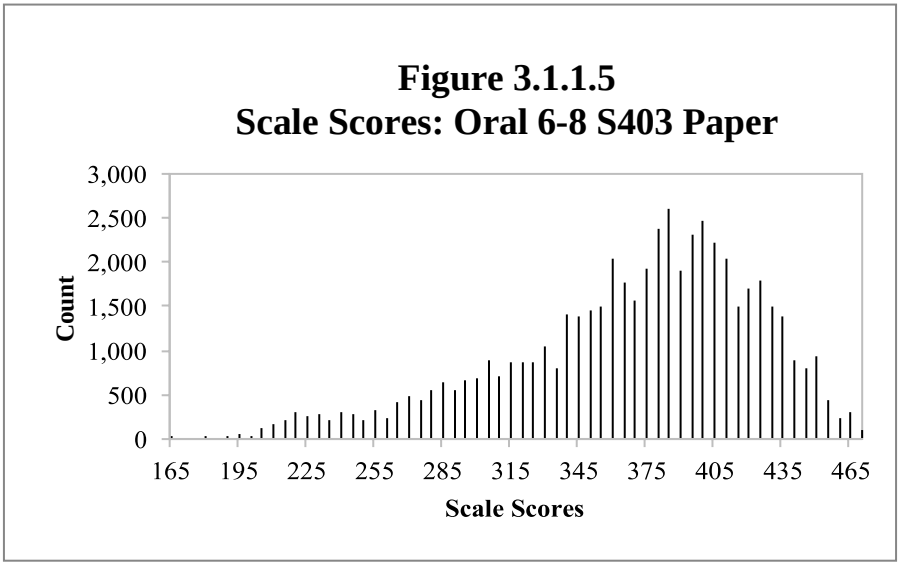
3.1.1.5 Grades 6–8

Table 3.1.1.5

Scale Score Descriptive Statistics: Oral 6-8 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	19,082	180	474	366.75	53.10
7	18,237	165	474	370.47	55.96
8	16,568	191	474	372.86	58.31
Total	53,887	165	474	369.89	55.77

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

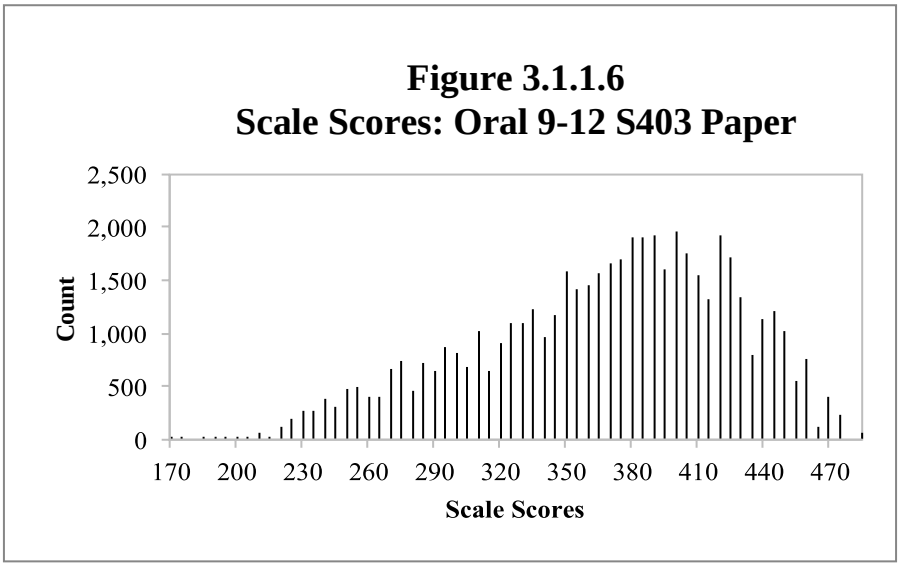
3.1.1.6 Grades 9–12

Table 3.1.1.6

Scale Score Descriptive Statistics: Oral 9-12 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	15,948	171	485	362.75	61.12
10	14,518	198	485	369.12	58.24
11	12,894	187	485	373.24	57.33
12	8,410	175	485	375.90	51.65
Total	51,770	171	485	369.29	58.12

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

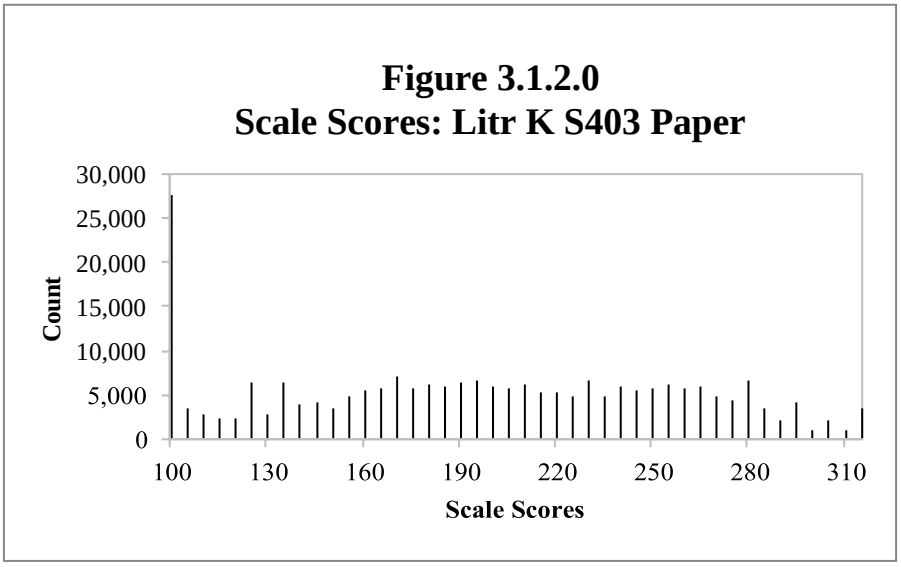
3.1.2 Literacy

3.1.2.0 Kindergarten

Table 3.1.2.0

Scale Score Descriptive Statistics: Litr K S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	232,449	100	315	196.50	62.33



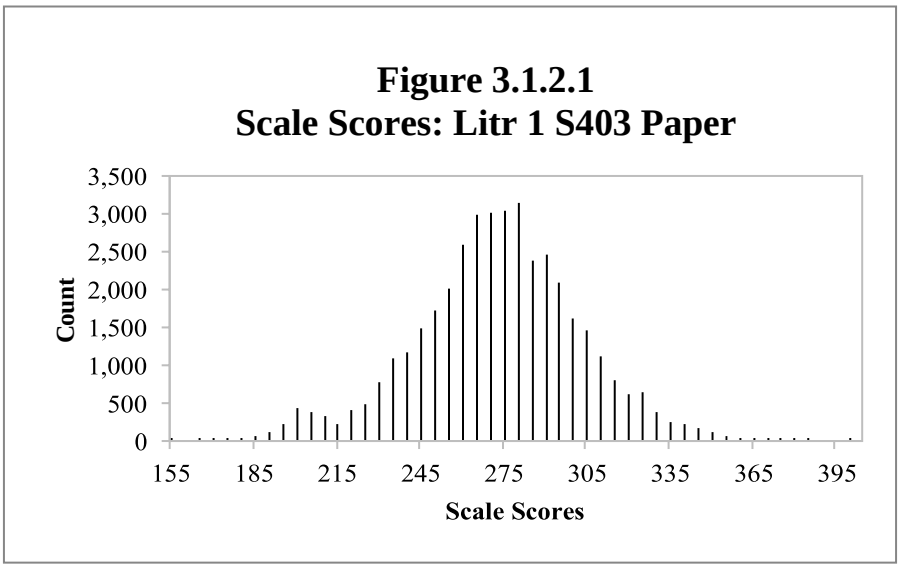
3.1.2.1 Grade 1

Table 3.1.2.1

Scale Score Descriptive Statistics: Litr 1 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	40,158	159	403	274.74	30.01

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

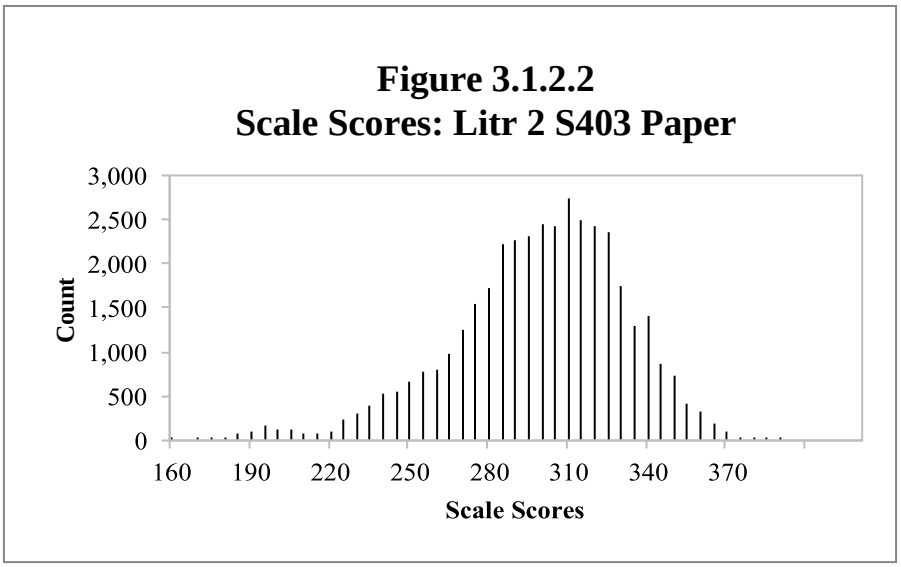
3.1.2.2 Grade 2

Table 3.1.2.2

Scale Score Descriptive Statistics: Litr 2 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	39,308	160	392	300.98	32.63

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

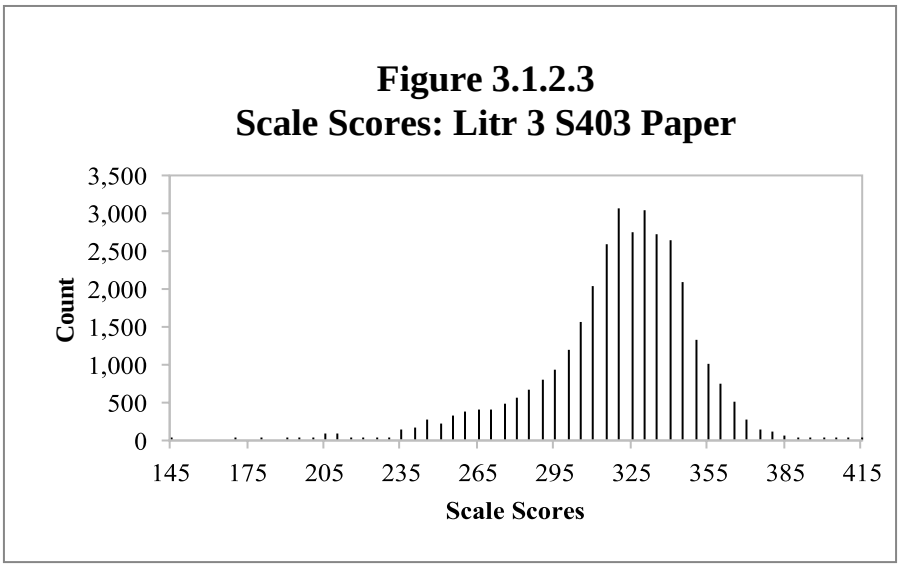
3.1.2.3 Grade 3

Table 3.1.2.3

Scale Score Descriptive Statistics: Litr 3 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	34,090	146	418	321.18	29.63

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

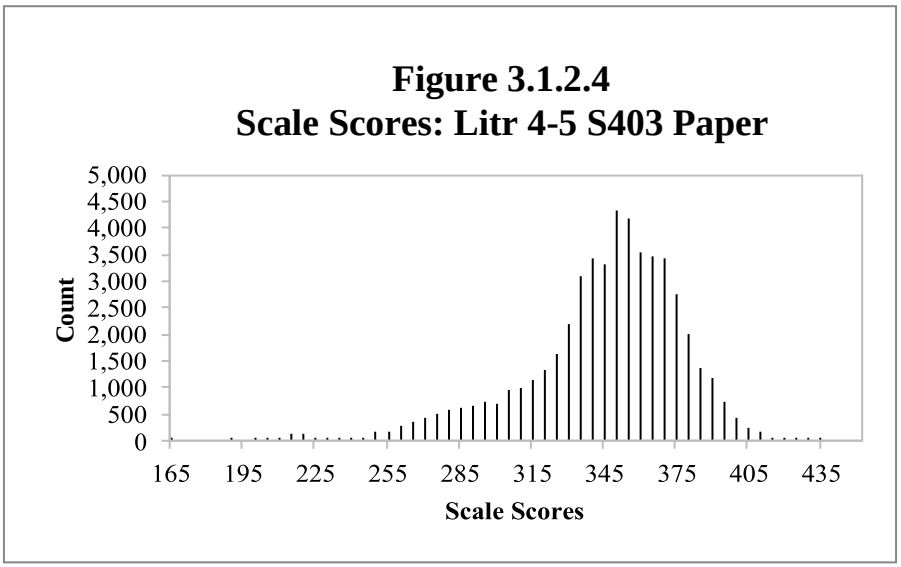
3.1.2.4 Grades 4–5

Table 3.1.2.4

Scale Score Descriptive Statistics: Litr 4-5 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	27,334	165	431	342.03	31.55
5	24,388	194	435	351.67	33.30
Total	51,722	165	435	346.57	32.75

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

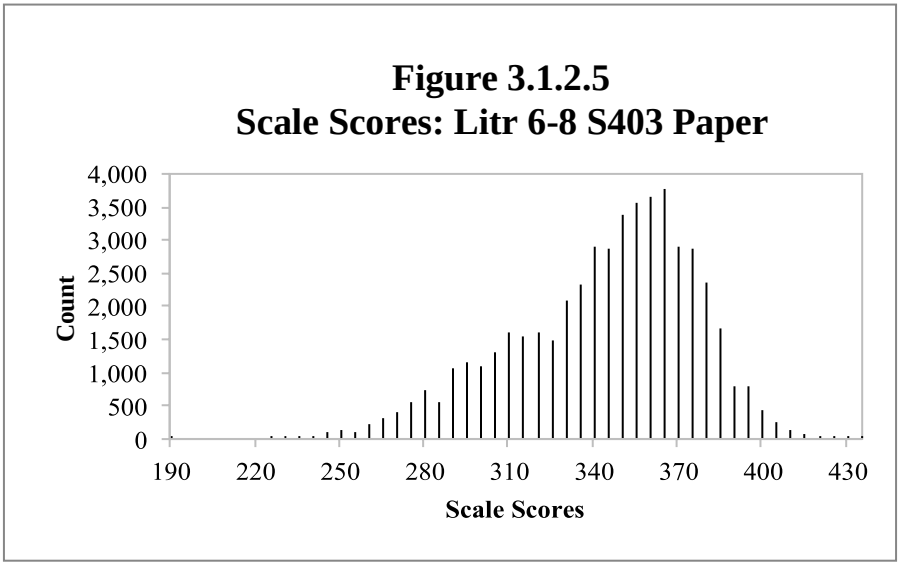
3.1.2.5 Grades 6–8

Table 3.1.2.5

Scale Score Descriptive Statistics: Litr 6-8 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,799	194	424	340.70	30.13
7	17,232	233	435	346.23	32.32
8	15,763	227	437	350.77	33.84
Total	50,794	194	437	345.70	32.32

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

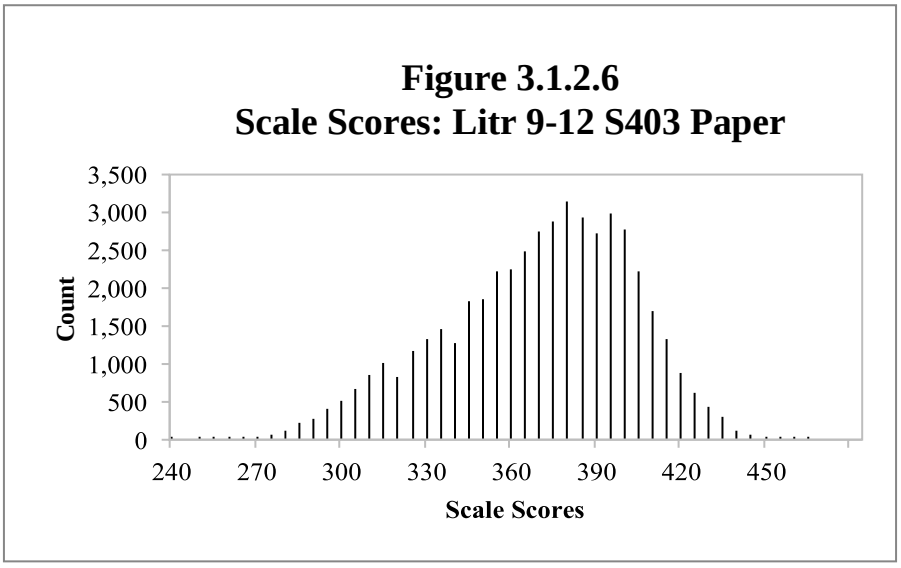
3.1.2.6 Grades 9–12

Table 3.1.2.6

Scale Score Descriptive Statistics: Litr 9-12 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	14,855	252	467	365.07	34.34
10	13,600	242	463	371.47	33.13
11	12,144	264	467	376.35	33.18
12	8,052	268	450	375.86	29.83
Total	48,651	242	467	371.46	33.33

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

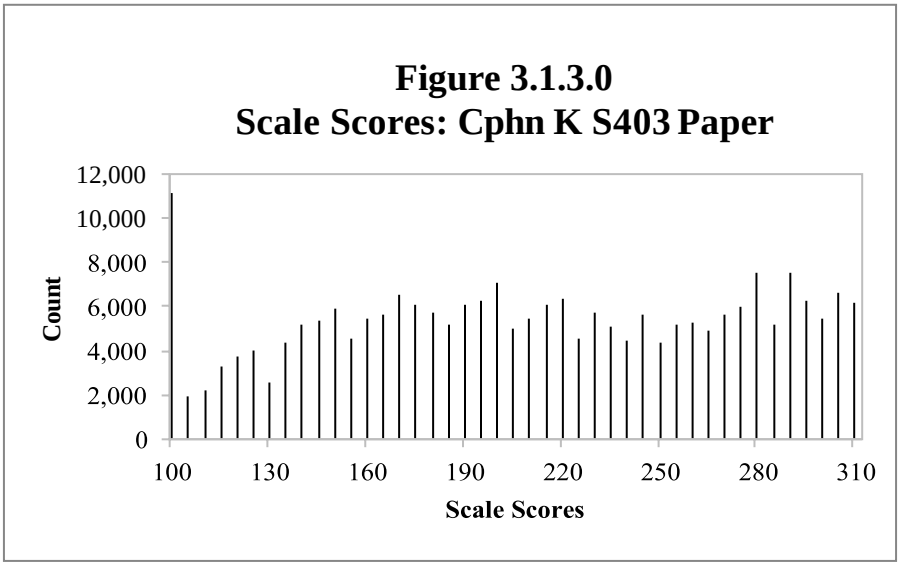
3.1.3 Comprehension

3.1.3.0 Kindergarten

Table 3.1.3.0

Scale Score Descriptive Statistics: Cphn K S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	232,455	100	312	212.28	61.62



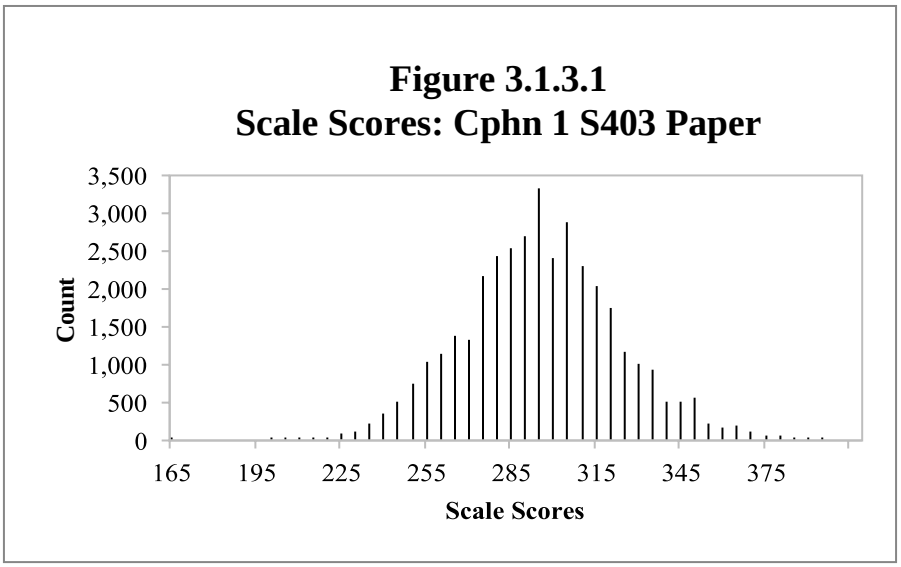
3.1.3.1 Grade 1

Table 3.1.3.1

Scale Score Descriptive Statistics: Cphn 1 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	37,051	167	397	297.74	27.34

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

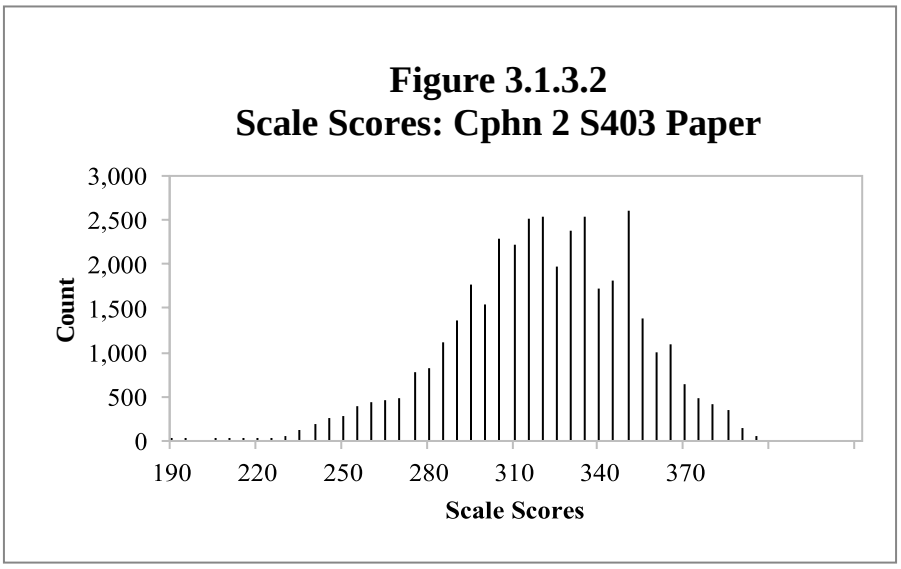
3.1.3.2 Grade 2

Table 3.1.3.2

Scale Score Descriptive Statistics: Cphn 2 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	38,279	193	397	322.68	31.33

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

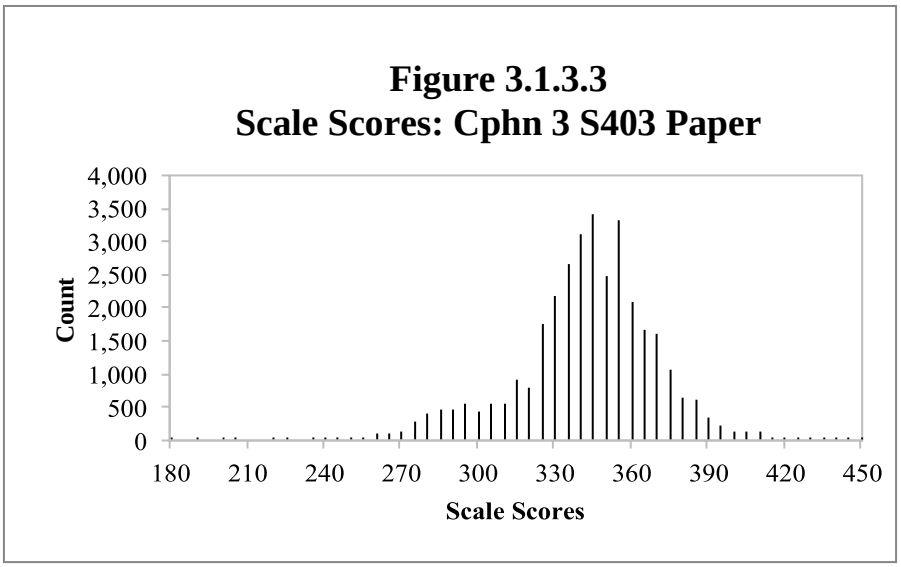
3.1.3.3 Grade 3

Table 3.1.3.3

Scale Score Descriptive Statistics: Cphn 3 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	33,502	181	453	344.36	26.84

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

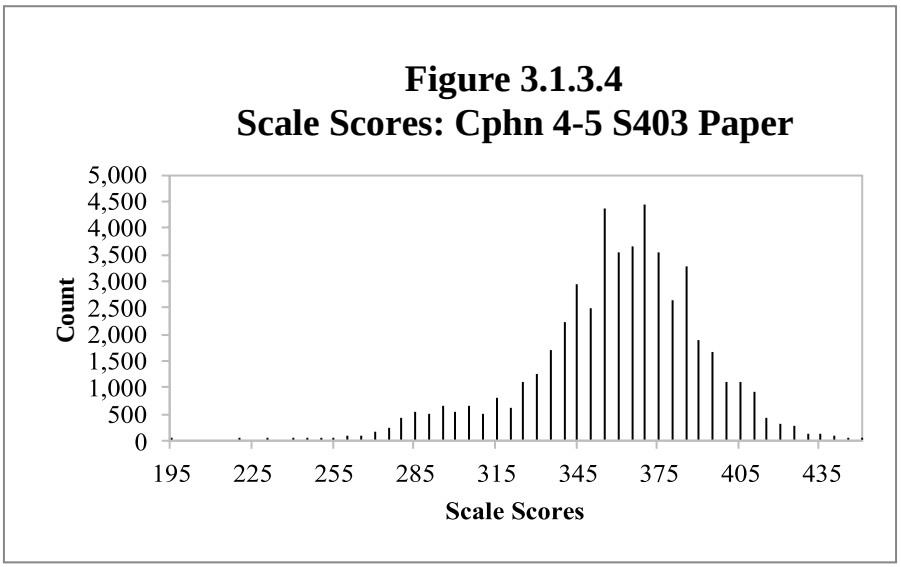
3.1.3.4 Grades 4–5

Table 3.1.3.4

Scale Score Descriptive Statistics: Cphn 4-5 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	26,980	220	453	357.19	29.63
5	24,146	199	453	366.32	32.08
Total	51,126	199	453	361.50	31.15

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

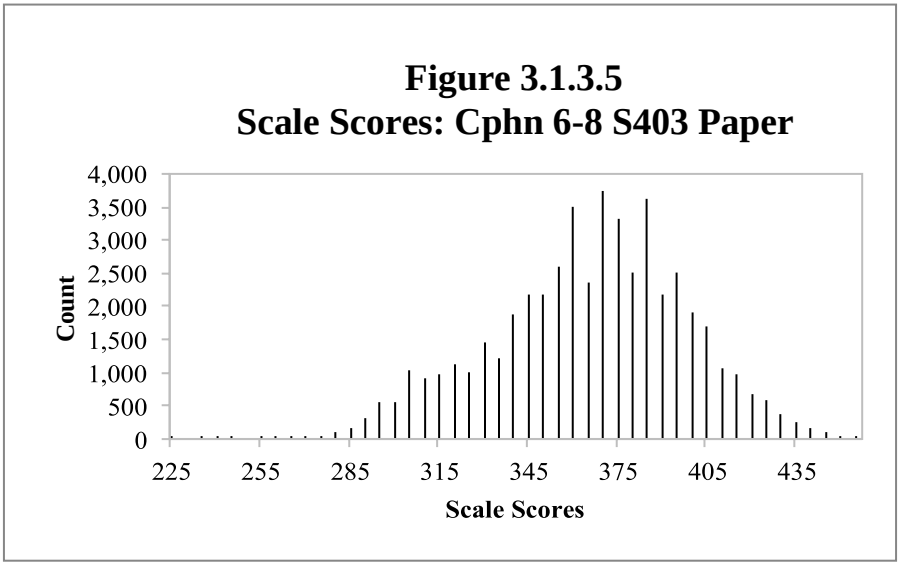
3.1.3.5 Grades 6–8

Table 3.1.3.5

Scale Score Descriptive Statistics: Cphn 6-8 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,481	229	459	362.11	29.11
7	16,914	240	459	368.13	32.47
8	15,463	239	459	372.98	34.73
Total	49,858	229	459	367.52	32.38

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

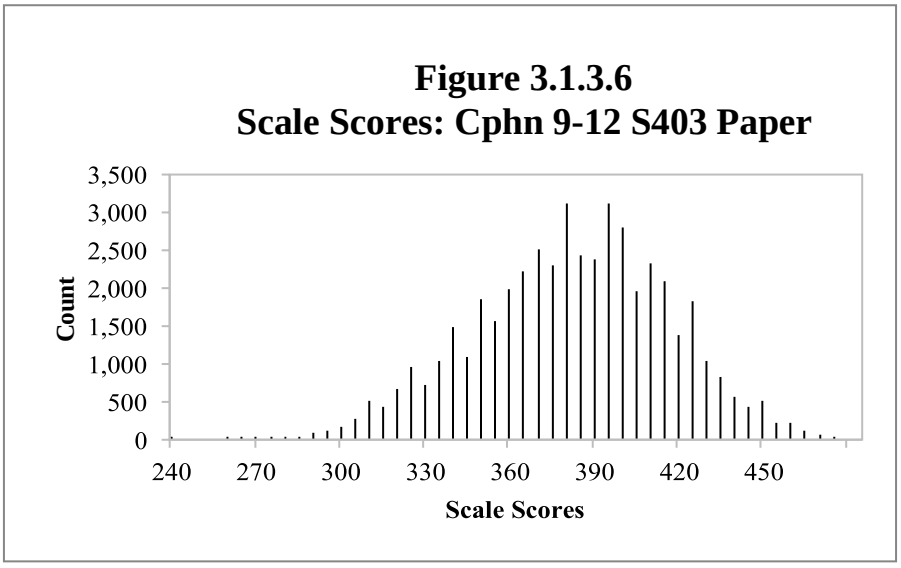
3.1.3.6 Grades 9–12

Table 3.1.3.6

Scale Score Descriptive Statistics: Cphn 9-12 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	14,463	266	477	378.62	34.66
10	13,269	243	477	384.79	34.11
11	11,881	263	477	388.76	34.80
12	7,832	267	477	388.04	30.82
Total	47,445	243	477	384.44	34.19

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

3.1.4 Overall

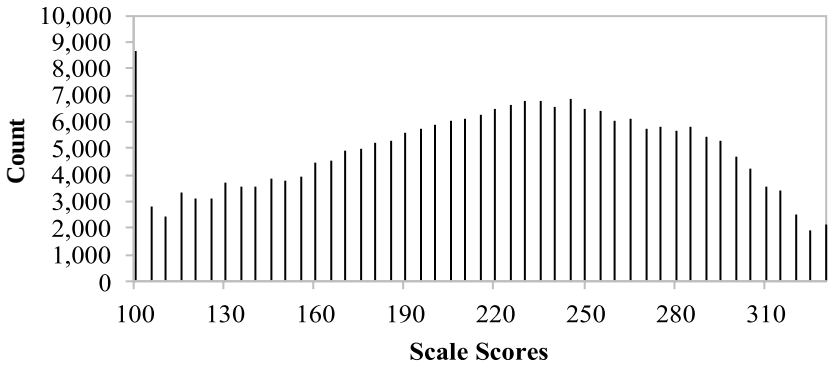
3.1.4.0 Kindergarten

Table 3.1.4.0

Scale Score Descriptive Statistics: Over K S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
K	232,443	100	333	218.94	61.32

Figure 3.1.4.0
Scale Scores: Over K S403 Paper



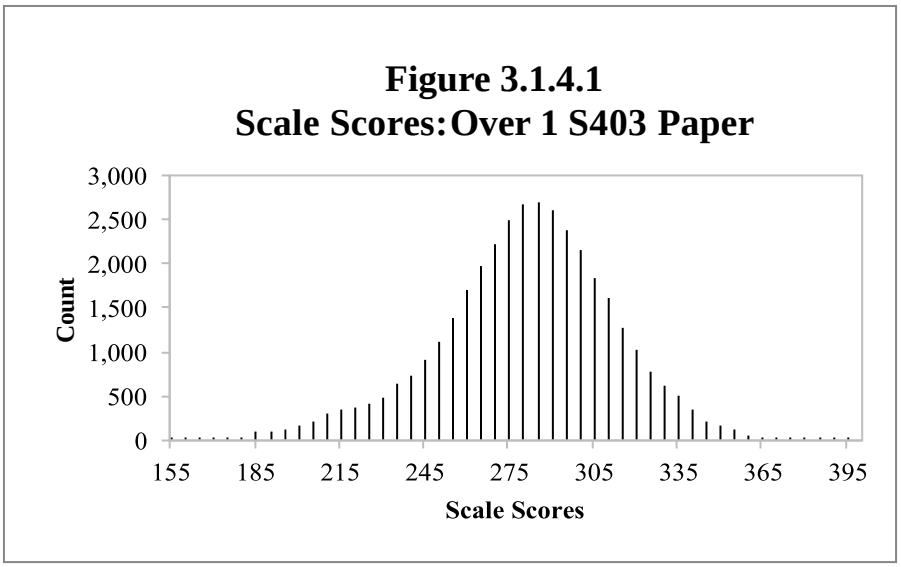
3.1.4.1 Grade 1

Table 3.1.4.1

Scale Score Descriptive Statistics: Over 1 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	36,884	158	398	282.64	31.09

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

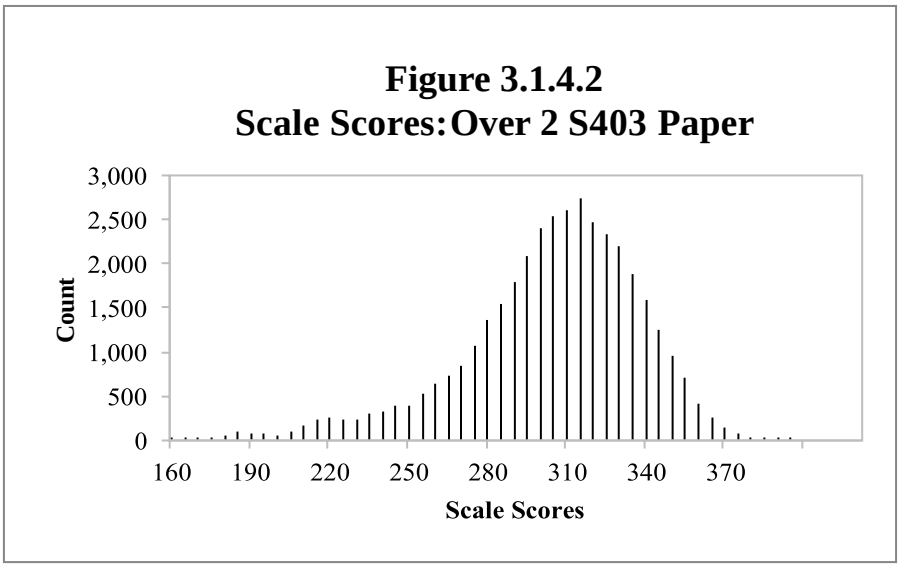
3.1.4.2 Grade 2

Table 3.1.4.2

Scale Score Descriptive Statistics: Over 2 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	38,092	162	397	306.26	33.60

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

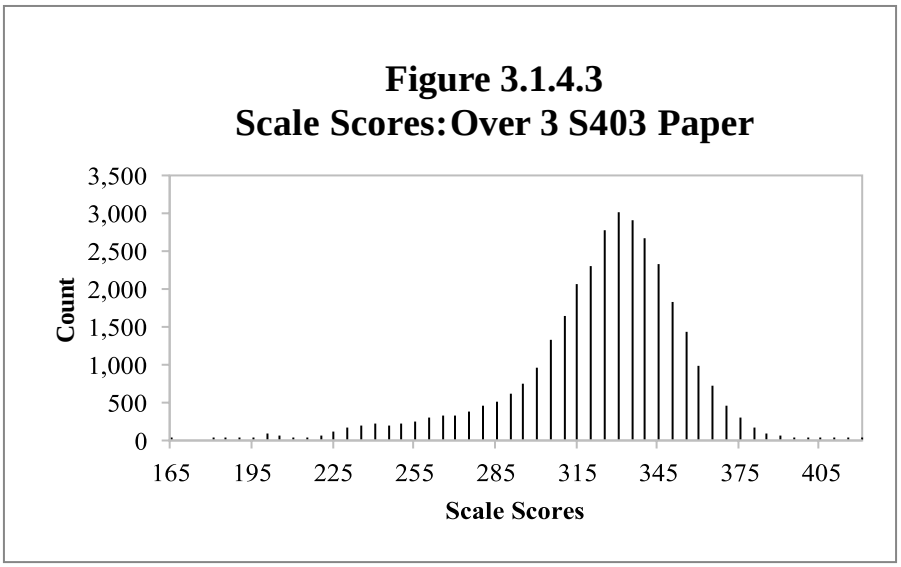
3.1.4.3 Grade 3

Table 3.1.4.3

Scale Score Descriptive Statistics: Over 3 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	33,333	167	423	324.74	31.42

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

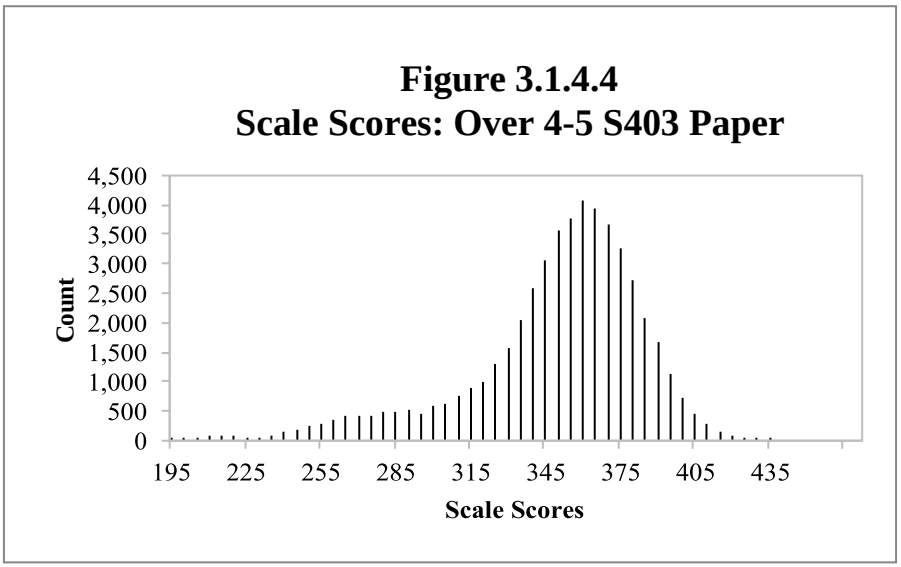
3.1.4.4 Grades 4–5

Table 3.1.4.4

Scale Score Descriptive Statistics: Over 4-5 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	26,877	195	434	346.81	34.11
5	24,072	196	435	356.02	35.99
Total	50,949	195	435	351.16	35.31

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

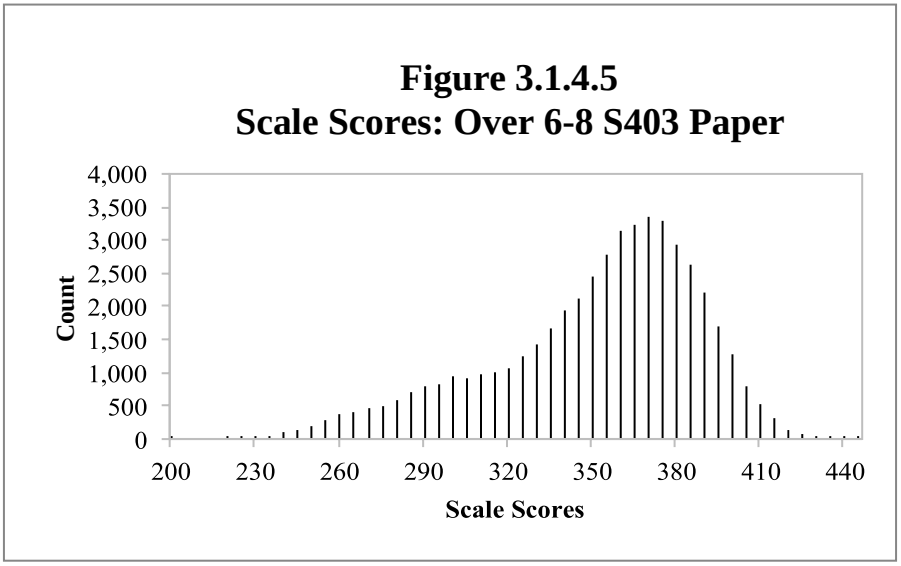
3.1.4.5 Grades 6–8

Table 3.1.4.5

Scale Score Descriptive Statistics: Over 6-8 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,393	203	433	348.60	34.64
7	16,839	223	445	353.82	37.18
8	15,385	223	448	357.73	39.07
Total	49,617	203	448	353.20	37.11

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

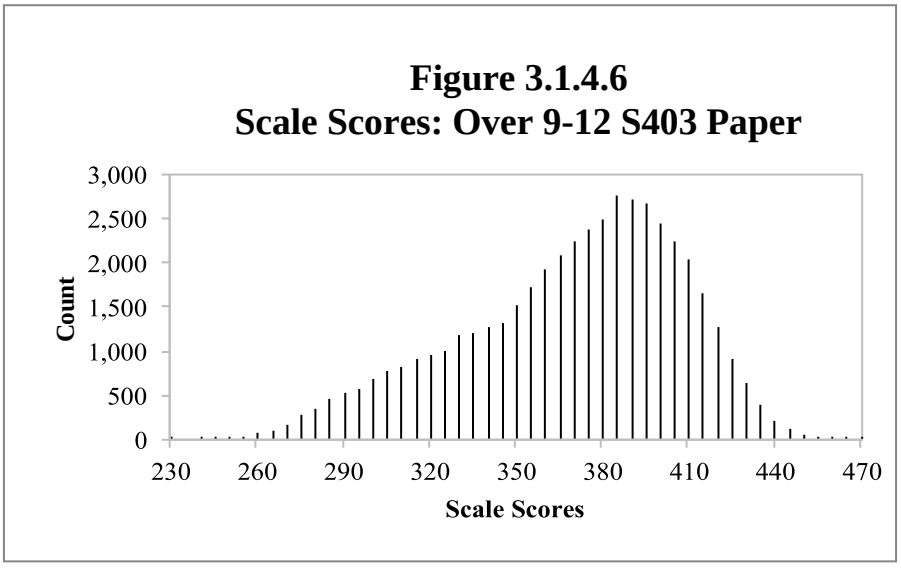
3.1.4.6 Grades 9–12

Table 3.1.4.6

Scale Score Descriptive Statistics: Over 9-12 S403 Paper

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	14,376	240	467	364.61	40.20
10	13,186	232	467	370.91	38.50
11	11,793	249	470	375.61	38.22
12	7,778	259	460	376.14	33.61
Total	47,133	232	470	371.03	38.49

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

3.2 Proficiency Level Distribution for Composites

Figures and tables in this section provide information on the proficiency level distribution for each of the composites for each grade-level cluster.

In each figure, the horizontal axis shows the six WIDA proficiency levels. The vertical axis shows the percentage of students. Each bar shows the percentage of students who were placed into each proficiency level in the domain being tested on this test form.

The tables in this section present, by grade and by total for the grade-level cluster:

- The WIDA proficiency level designation (1–6)
- The number of students (count) whose performance on the test form placed them into that proficiency level in the domain being tested
- The percentage of students, out of the total number of students taking the form, who were placed into that proficiency level in the domain being tested

In the tables and figures in this section, scale scores which were computed using mode-adjusted scoring tables are excluded from the analysis.

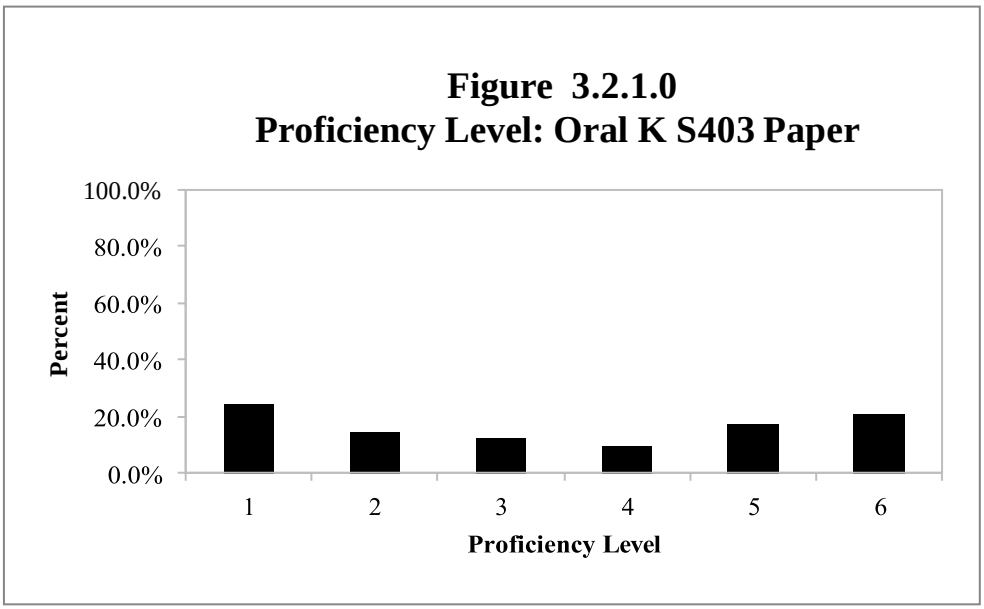
3.2.1 Oral

3.2.1.0 Kindergarten

Table 3.2.1.0

Proficiency Level Distribution: Oral K S403 Paper

Level	Count	Percent
1	57,054	24.5%
2	33,940	14.6%
3	29,154	12.5%
4	22,549	9.7%
5	40,332	17.3%
6	49,433	21.3%
Total	232,462	100.0%



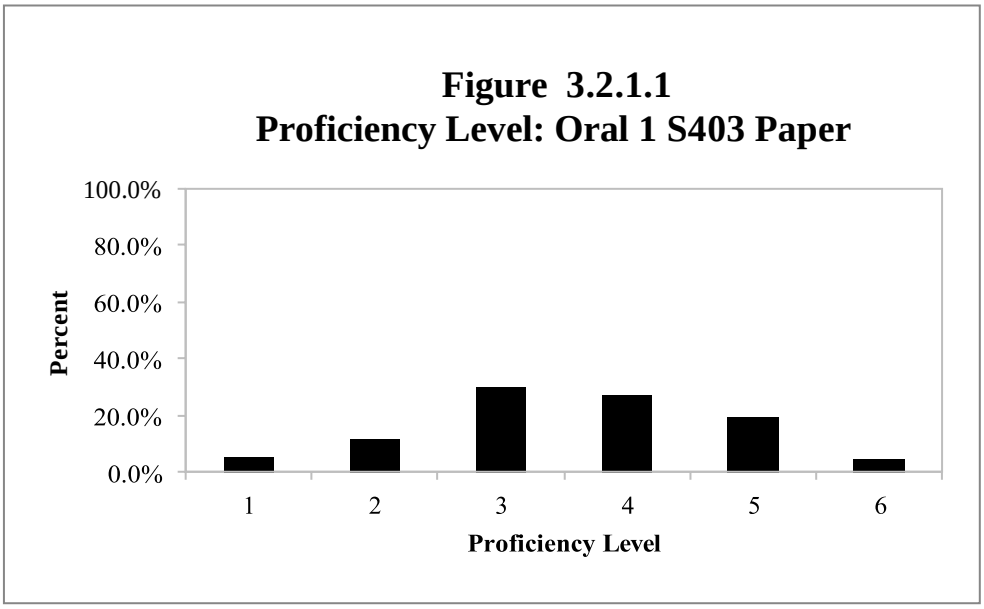
3.2.1.1 Grade 1

Table 3.2.1.1

Proficiency Level Distribution: Oral 1 S403 Paper

Level	Count	Percent
1	2,539	5.9%
2	5,205	12.0%
3	13,131	30.3%
4	11,971	27.6%
5	8,409	19.4%
6	2,088	4.8%
Total	43,343	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

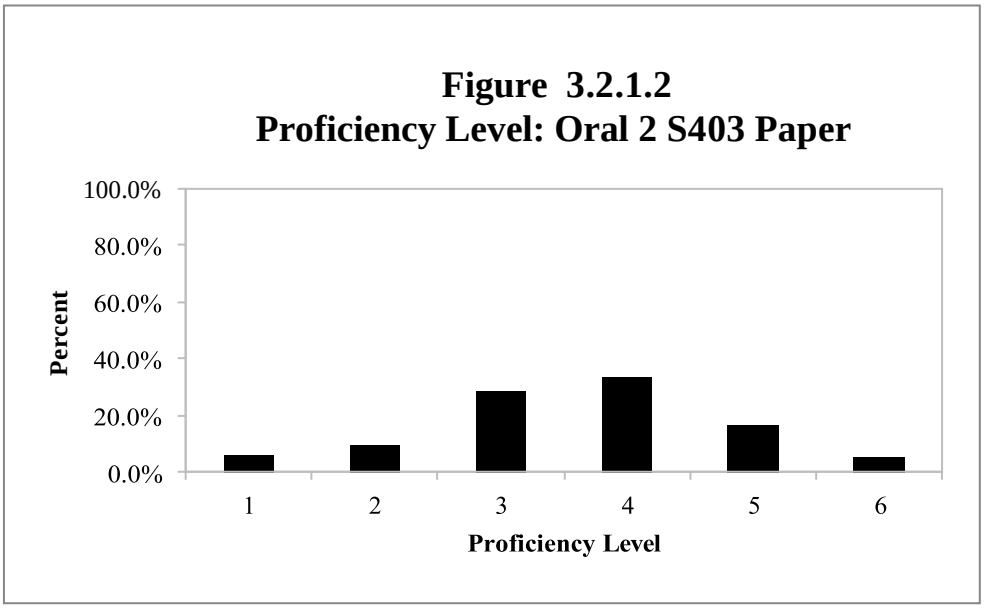
3.2.1.2 Grade 2

Table 3.2.1.2

Proficiency Level Distribution: Oral 2 S403 Paper

Level	Count	Percent
1	2,595	6.1%
2	4,167	9.7%
3	12,238	28.6%
4	14,486	33.8%
5	7,041	16.4%
6	2,329	5.4%
Total	42,856	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

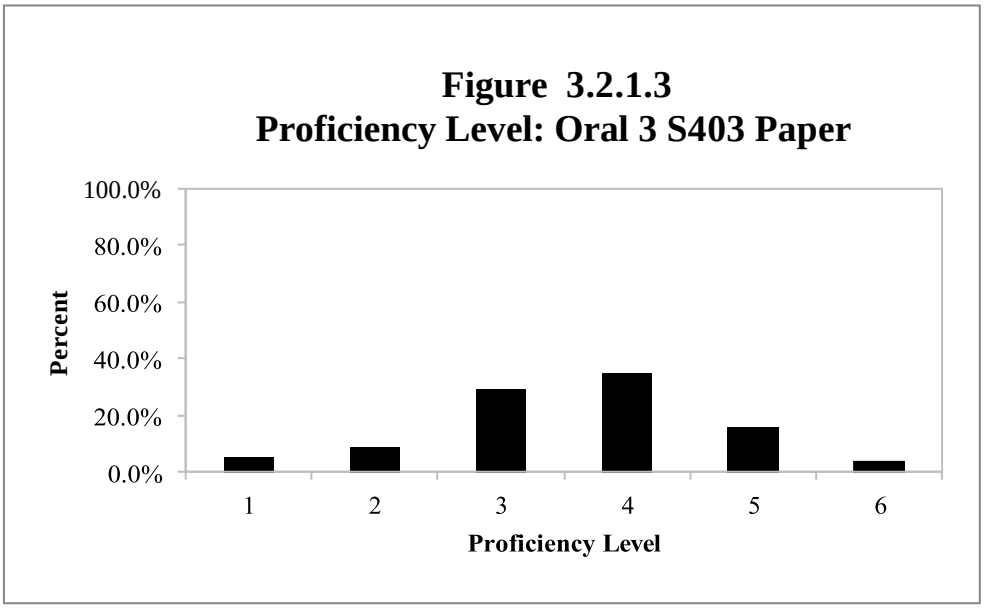
3.2.1.3 Grade 3

Table 3.2.1.3

Proficiency Level Distribution: Oral 3 S403 Paper

Level	Count	Percent
1	2,084	5.5%
2	3,404	9.1%
3	11,088	29.5%
4	13,202	35.1%
5	6,164	16.4%
6	1,652	4.4%
Total	37,594	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

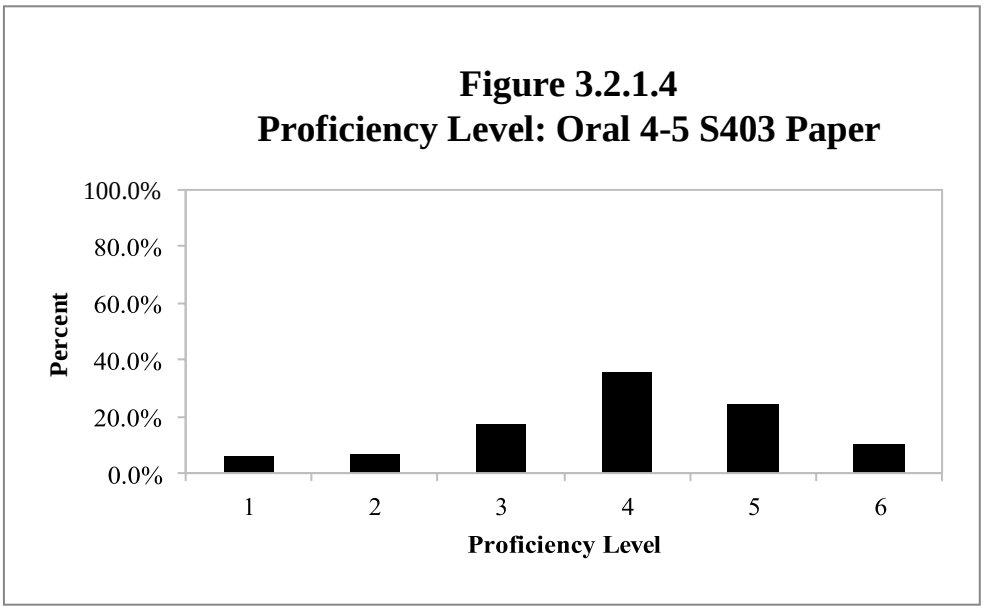
3.2.1.4 Grades 4–5

Table 3.2.1.4

Proficiency Level Distribution: Oral 4-5 S403 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	1,701	5.7%	1,718	6.6%	3,419	6.1%
2	1,955	6.6%	1,721	6.6%	3,676	6.6%
3	5,331	18.0%	4,315	16.5%	9,646	17.3%
4	10,281	34.7%	9,566	36.5%	19,847	35.6%
5	7,175	24.2%	6,386	24.4%	13,561	24.3%
6	3,164	10.7%	2,488	9.5%	5,652	10.1%
Total	29,607	100.0%	26,194	100.0%	55,801	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

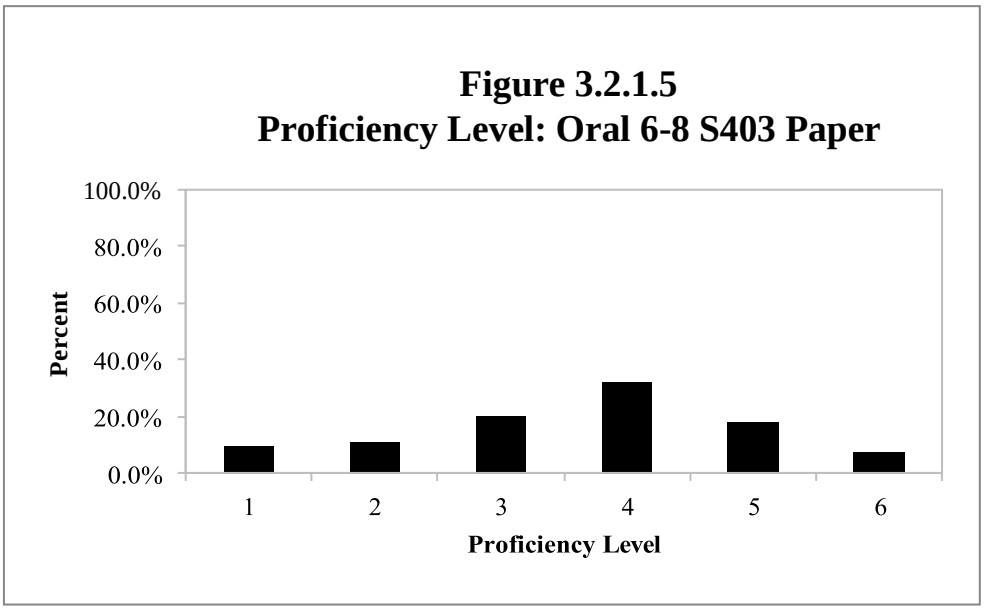
3.2.1.5 Grades 6–8

Table 3.2.1.5

Proficiency Level Distribution: Oral 6-8 S403 Paper

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,528	8.0%	1,875	10.3%	2,009	12.1%	5,412	10.0%
2	1,857	9.7%	1,978	10.8%	2,002	12.1%	5,837	10.8%
3	3,911	20.5%	3,778	20.7%	3,386	20.4%	11,075	20.6%
4	6,294	33.0%	5,924	32.5%	5,148	31.1%	17,366	32.2%
5	3,841	20.1%	3,288	18.0%	2,776	16.8%	9,905	18.4%
6	1,651	8.7%	1,394	7.6%	1,247	7.5%	4,292	8.0%
Total	19,082	100.0%	18,237	100.0%	16,568	100.0%	53,887	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

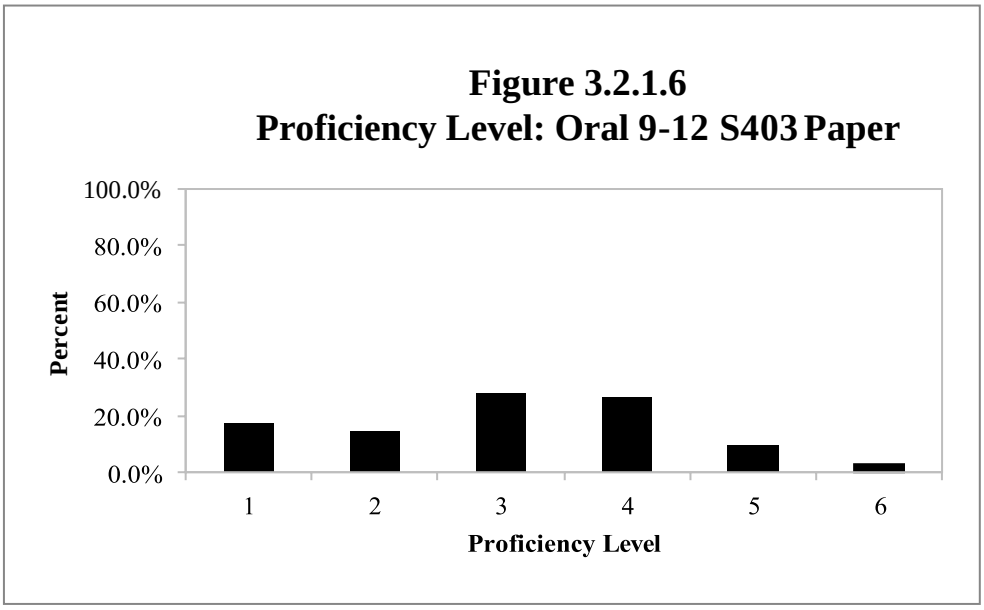
3.2.1.6 Grades 9–12

Table 3.2.1.6

Proficiency Level Distribution: Oral 9-12 S403 Paper

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	2,951	18.5%	2,589	17.8%	2,301	17.8%	1,348	16.0%	9,189	17.7%
2	2,292	14.4%	2,039	14.0%	1,818	14.1%	1,260	15.0%	7,409	14.3%
3	3,865	24.2%	3,966	27.3%	3,680	28.5%	2,899	34.5%	14,410	27.8%
4	4,293	26.9%	3,853	26.5%	3,389	26.3%	2,168	25.8%	13,703	26.5%
5	1,920	12.0%	1,449	10.0%	1,265	9.8%	520	6.2%	5,154	10.0%
6	627	3.9%	622	4.3%	441	3.4%	215	2.6%	1,905	3.7%
Total	15,948	100.0%	14,518	100.0%	12,894	100.0%	8,410	100.0%	51,770	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

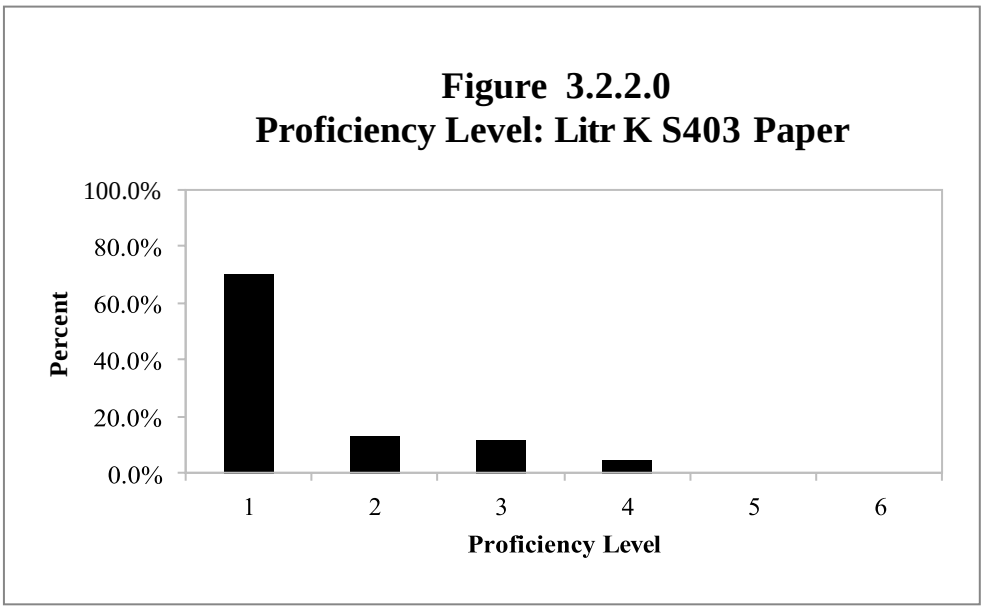
3.2.2 Literacy

3.2.2.0 Kindergarten

Table 3.2.2.0

Proficiency Level Distribution: Litr K S403 Paper

Level	Count	Percent
1	162,486	69.9%
2	30,687	13.2%
3	27,328	11.8%
4	11,948	5.1%
5	0	0.0%
6	0	0.0%
Total	232,449	100.0%



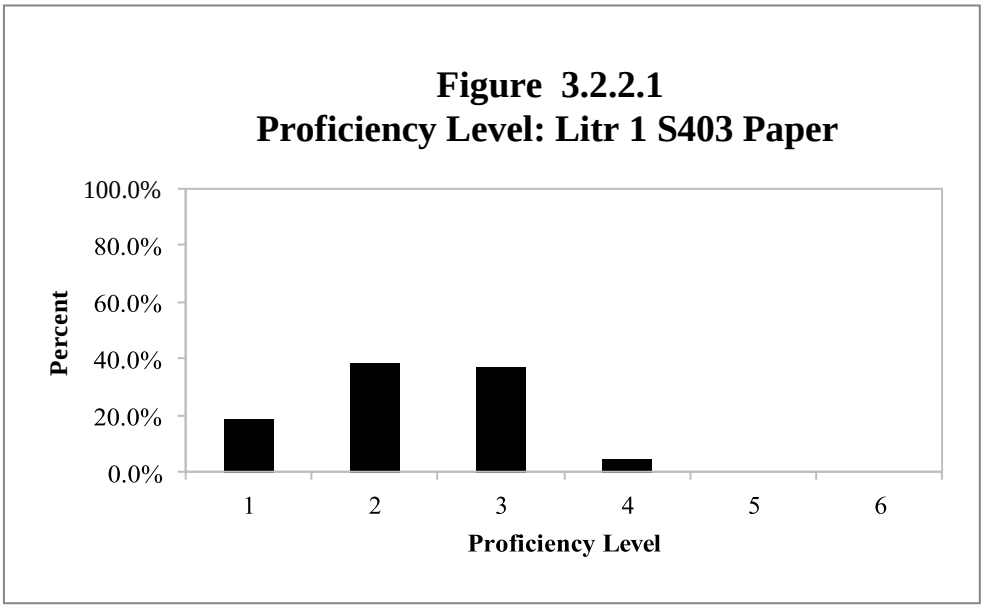
3.2.2.1 Grade 1

Table 3.2.2.1

Proficiency Level Distribution: Litr 1 S403 Paper

Level	Count	Percent
1	7,599	18.9%
2	15,334	38.2%
3	14,855	37.0%
4	2,073	5.2%
5	271	0.7%
6	26	0.1%
Total	40,158	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

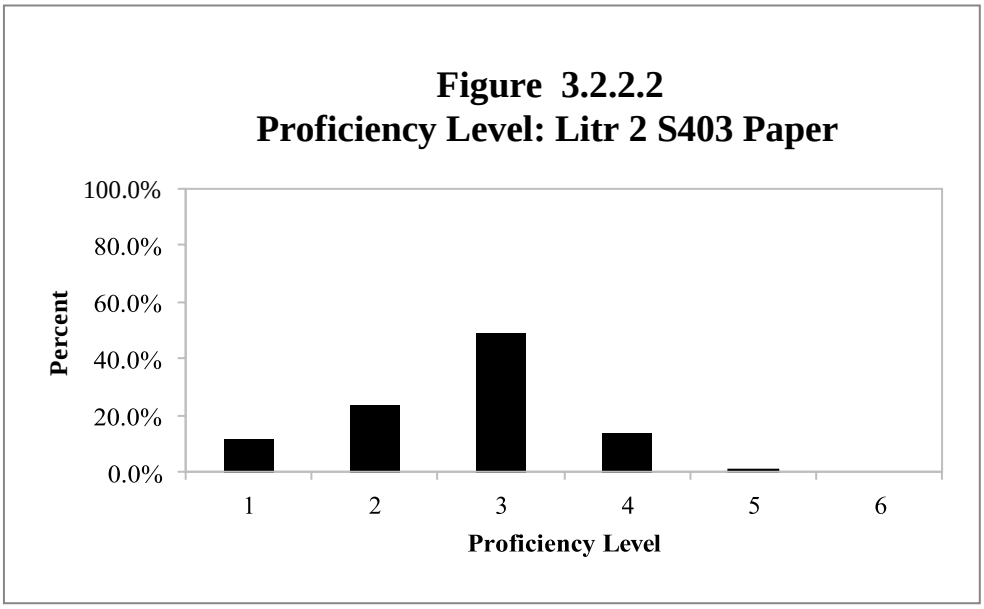
3.2.2.2 Grade 2

Table 3.2.2.2

Proficiency Level Distribution: Litr 2 S403 Paper

Level	Count	Percent
1	4,748	12.1%
2	9,421	24.0%
3	19,158	48.7%
4	5,526	14.1%
5	440	1.1%
6	15	0.0%
Total	39,308	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

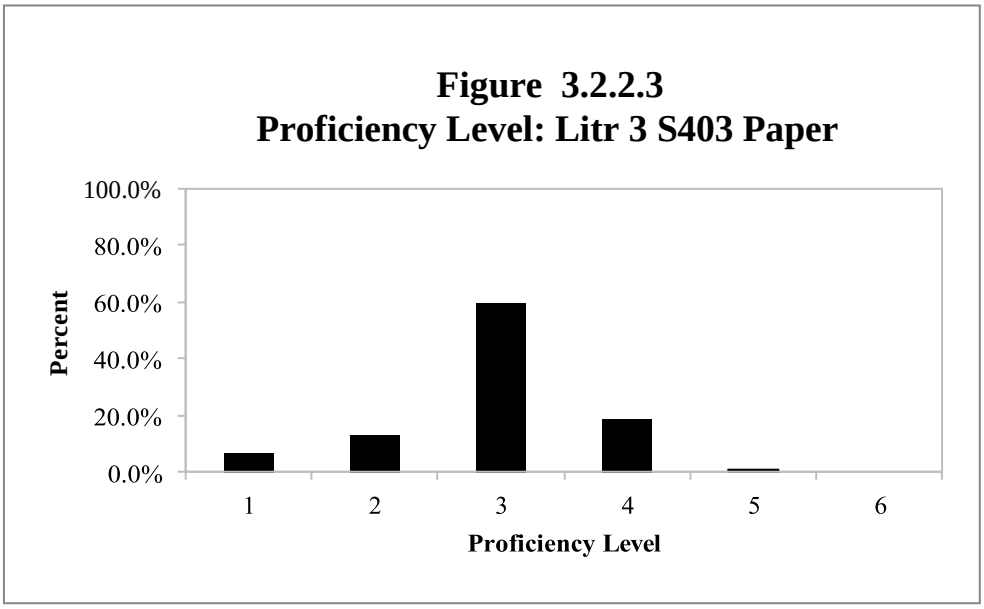
3.2.2.3 Grade 3

Table 3.2.2.3

Proficiency Level Distribution: Litr 3 S403 Paper

Level	Count	Percent
1	2,475	7.3%
2	4,413	12.9%
3	20,318	59.6%
4	6,385	18.7%
5	452	1.3%
6	47	0.1%
Total	34,090	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

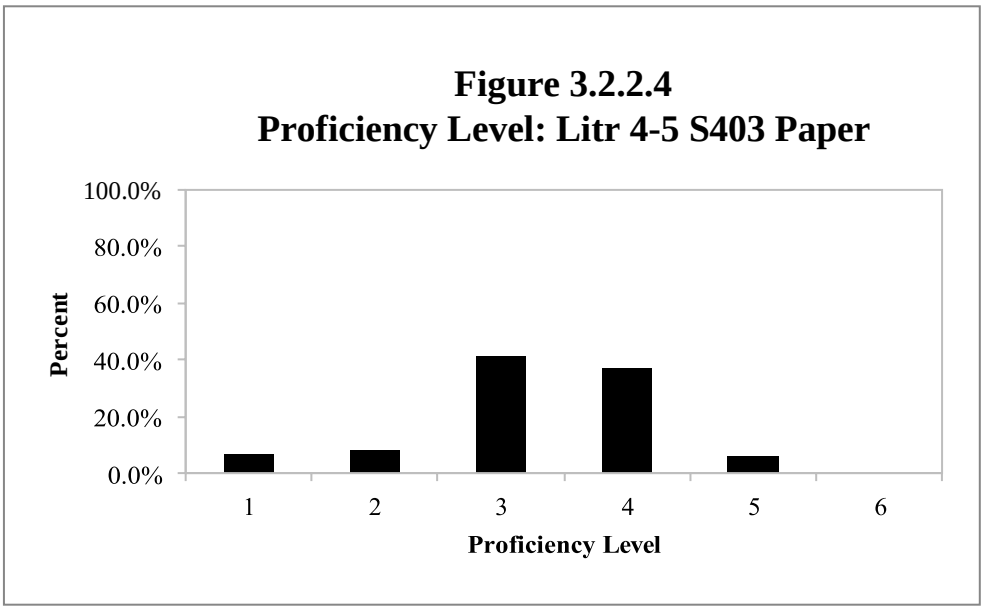
3.2.2.4 Grades 4–5

Table 3.2.2.4

Proficiency Level Distribution: Litr 4-5 S403 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	1,888	6.9%	1,608	6.6%	3,496	6.8%
2	2,075	7.6%	2,060	8.4%	4,135	8.0%
3	12,114	44.3%	9,235	37.9%	21,349	41.3%
4	9,914	36.3%	9,409	38.6%	19,323	37.4%
5	1,180	4.3%	1,920	7.9%	3,100	6.0%
6	163	0.6%	156	0.6%	319	0.6%
Total	27,334	100.0%	24,388	100.0%	51,722	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

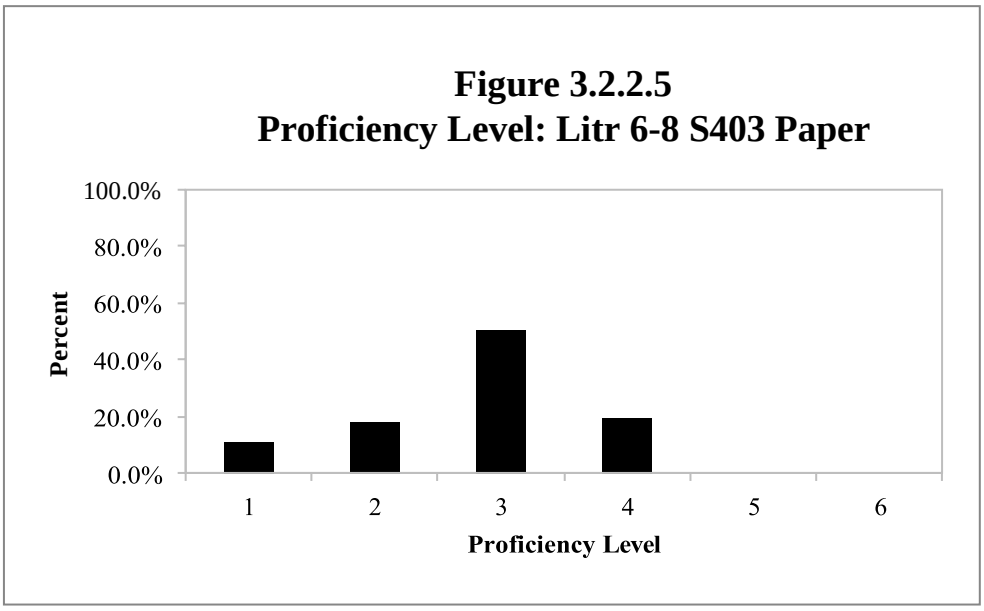
3.2.2.5 Grades 6–8

Table 3.2.2.5

Proficiency Level Distribution: Litr 6-8 S403 Paper

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,673	9.4%	1,865	10.8%	2,117	13.4%	5,655	11.1%
2	3,139	17.6%	3,152	18.3%	2,863	18.2%	9,154	18.0%
3	9,503	53.4%	8,800	51.1%	7,435	47.2%	25,738	50.7%
4	3,358	18.9%	3,243	18.8%	3,191	20.2%	9,792	19.3%
5	120	0.7%	165	1.0%	156	1.0%	441	0.9%
6	6	0.0%	7	0.0%	1	0.0%	14	0.0%
Total	17,799	100.0%	17,232	100.0%	15,763	100.0%	50,794	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

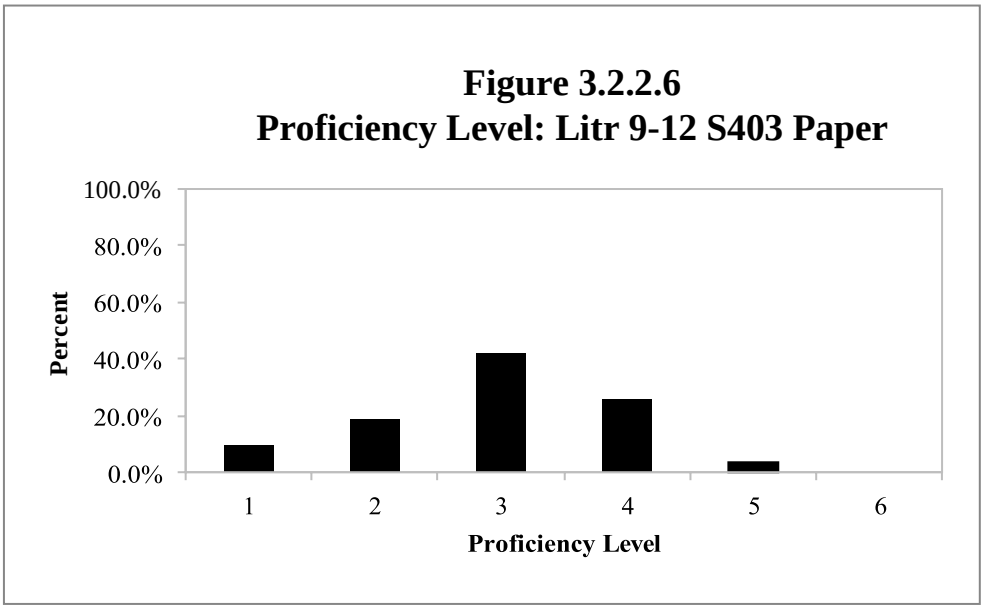
3.2.2.6 Grades 9–12

Table 3.2.2.6

Proficiency Level Distribution: Litr 9-12 S403 Paper

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,475	9.9%	1,199	8.8%	1,163	9.6%	826	10.3%	4,663	9.6%
2	2,628	17.7%	2,357	17.3%	2,321	19.1%	1,816	22.6%	9,122	18.7%
3	5,998	40.4%	5,719	42.1%	4,885	40.2%	3,838	47.7%	20,440	42.0%
4	4,048	27.3%	3,722	27.4%	3,266	26.9%	1,431	17.8%	12,467	25.6%
5	688	4.6%	595	4.4%	507	4.2%	141	1.8%	1,931	4.0%
6	18	0.1%	8	0.1%	2	0.0%	0	0.0%	28	0.1%
Total	14,855	100.0%	13,600	100.0%	12,144	100.0%	8,052	100.0%	48,651	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

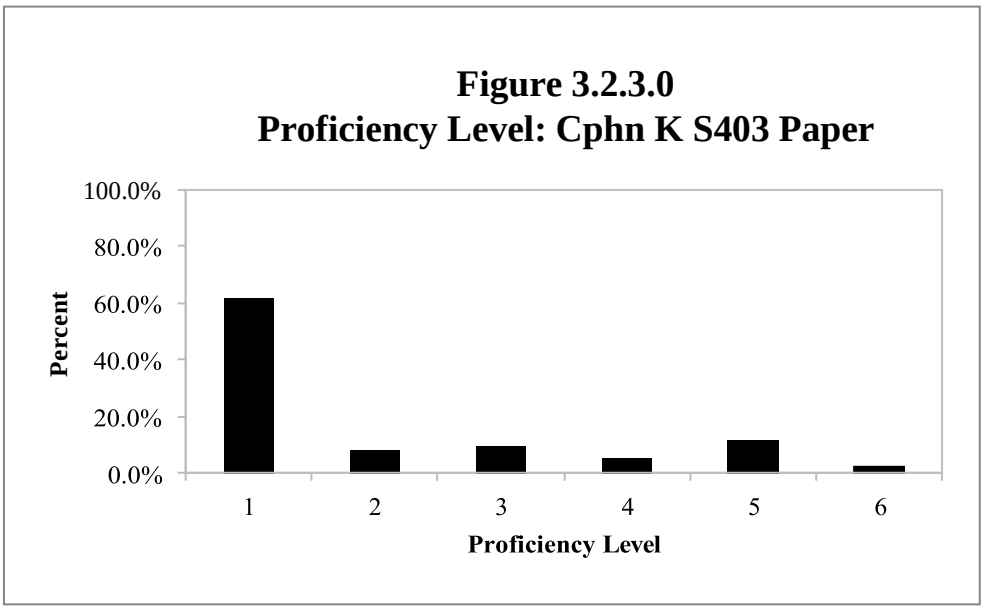
3.2.3 Comprehension

3.2.3.0 Kindergarten

Table 3.2.3.0

Proficiency Level Distribution: Cphn K S403 Paper

Level	Count	Percent
1	143,875	61.9%
2	19,781	8.5%
3	23,216	10.0%
4	12,109	5.2%
5	27,301	11.7%
6	6,173	2.7%
Total	232,455	100.0%



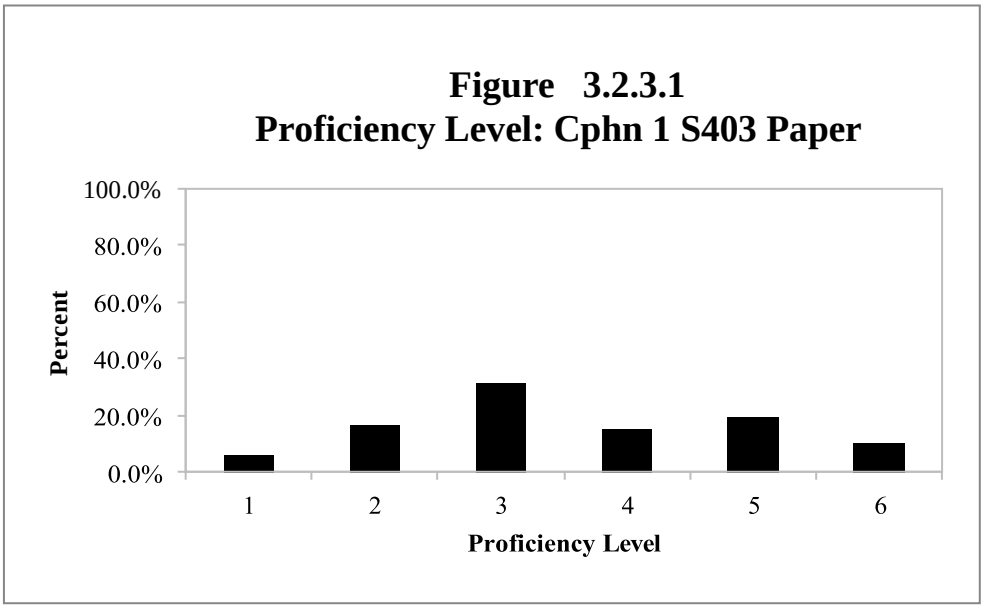
3.2.3.1 Grade 1

Table 3.2.3.1

Proficiency Level Distribution: Cphn 1 S403 Paper

Level	Count	Percent
1	2,245	6.1%
2	6,129	16.5%
3	11,729	31.7%
4	5,671	15.3%
5	7,316	19.7%
6	3,961	10.7%
Total	37,051	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

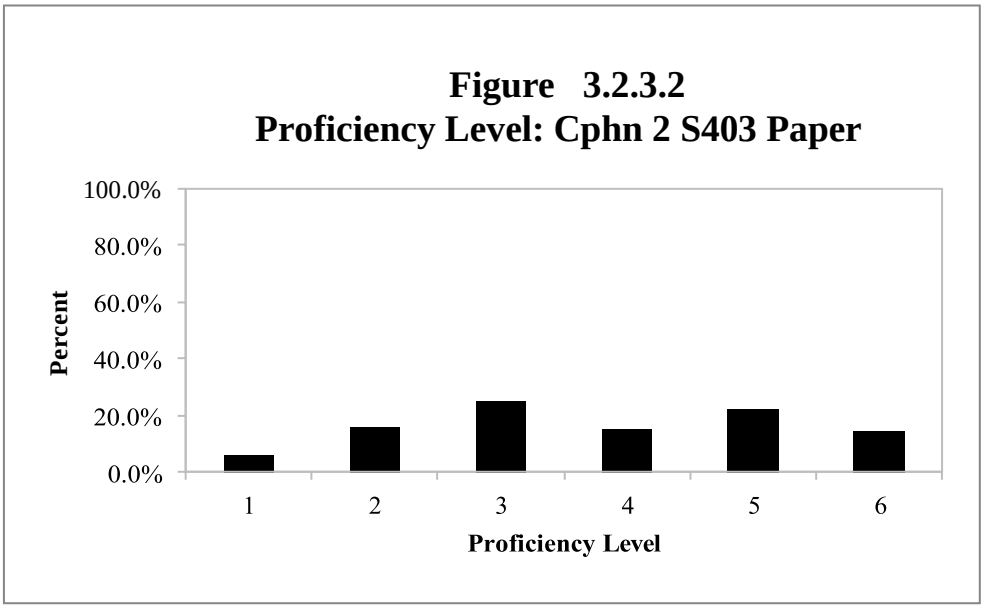
3.2.3.2 Grade 2

Table 3.2.3.2

Proficiency Level Distribution: Cphn 2 S403 Paper

Level	Count	Percent
1	2,392	6.2%
2	6,183	16.2%
3	9,564	25.0%
4	5,876	15.4%
5	8,679	22.7%
6	5,585	14.6%
Total	38,279	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

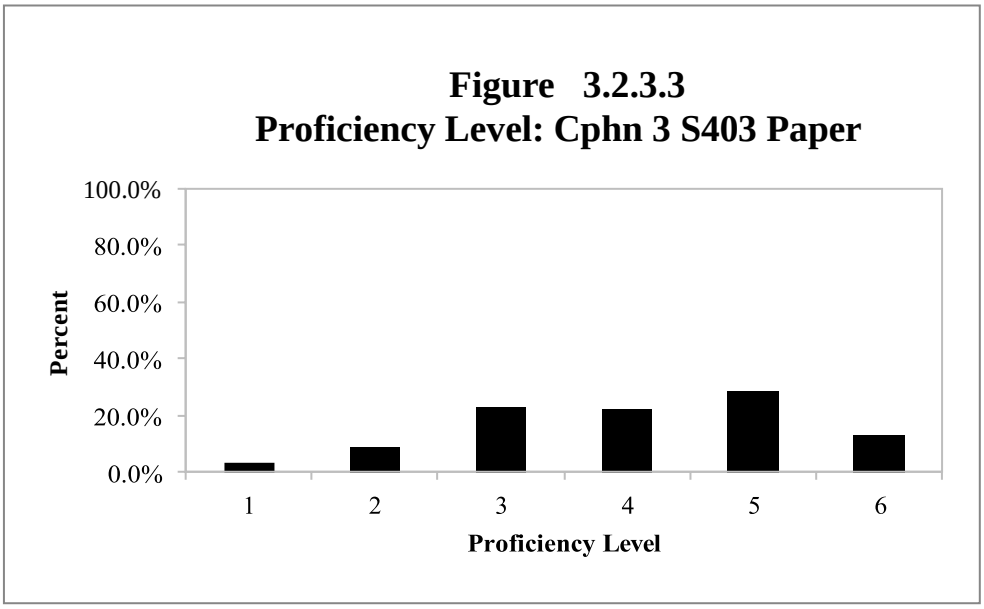
3.2.3.3 Grade 3

Table 3.2.3.3

Proficiency Level Distribution: Cphn 3 S403 Paper

Level	Count	Percent
1	1,219	3.6%
2	3,010	9.0%
3	7,637	22.8%
4	7,557	22.6%
5	9,693	28.9%
6	4,386	13.1%
Total	33,502	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

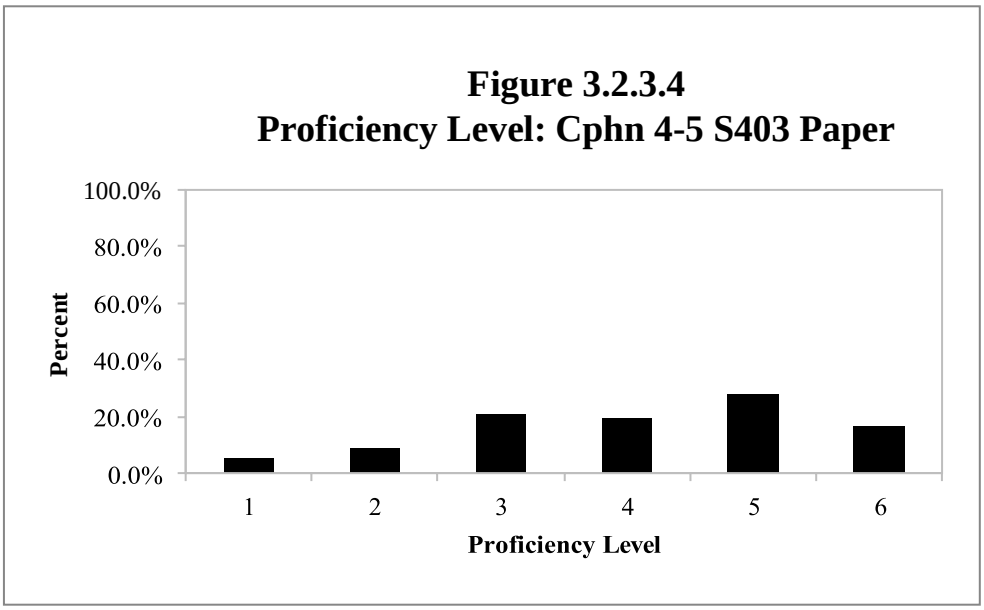
3.2.3.4 Grades 4–5

Table 3.2.3.4

Proficiency Level Distribution: Cphn 4-5 S403 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	1,387	5.1%	1,529	6.3%	2,916	5.7%
2	2,260	8.4%	2,282	9.5%	4,542	8.9%
3	5,731	21.2%	4,926	20.4%	10,657	20.8%
4	5,309	19.7%	4,574	18.9%	9,883	19.3%
5	8,067	29.9%	6,412	26.6%	14,479	28.3%
6	4,226	15.7%	4,423	18.3%	8,649	16.9%
Total	26,980	100.0%	24,146	100.0%	51,126	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

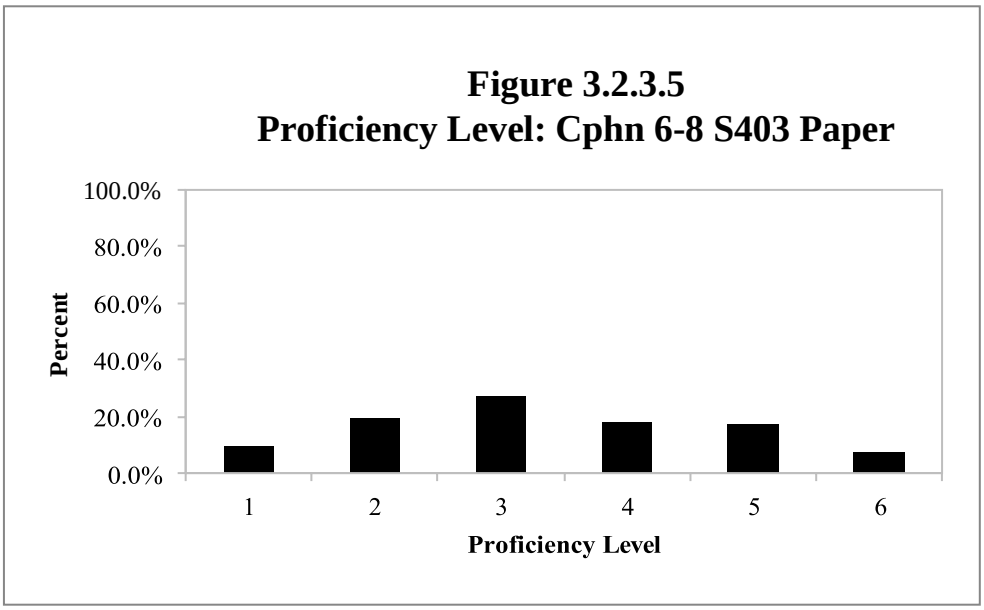
3.2.3.5 Grades 6–8

Table 3.2.3.5

Proficiency Level Distribution: Cphn 6-8 S403 Paper

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,253	7.2%	1,706	10.1%	1,854	12.0%	4,813	9.7%
2	3,350	19.2%	3,281	19.4%	3,128	20.2%	9,759	19.6%
3	5,247	30.0%	4,671	27.6%	3,835	24.8%	13,753	27.6%
4	3,460	19.8%	3,028	17.9%	2,418	15.6%	8,906	17.9%
5	2,953	16.9%	2,809	16.6%	2,981	19.3%	8,743	17.5%
6	1,218	7.0%	1,419	8.4%	1,247	8.1%	3,884	7.8%
Total	17,481	100.0%	16,914	100.0%	15,463	100.0%	49,858	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

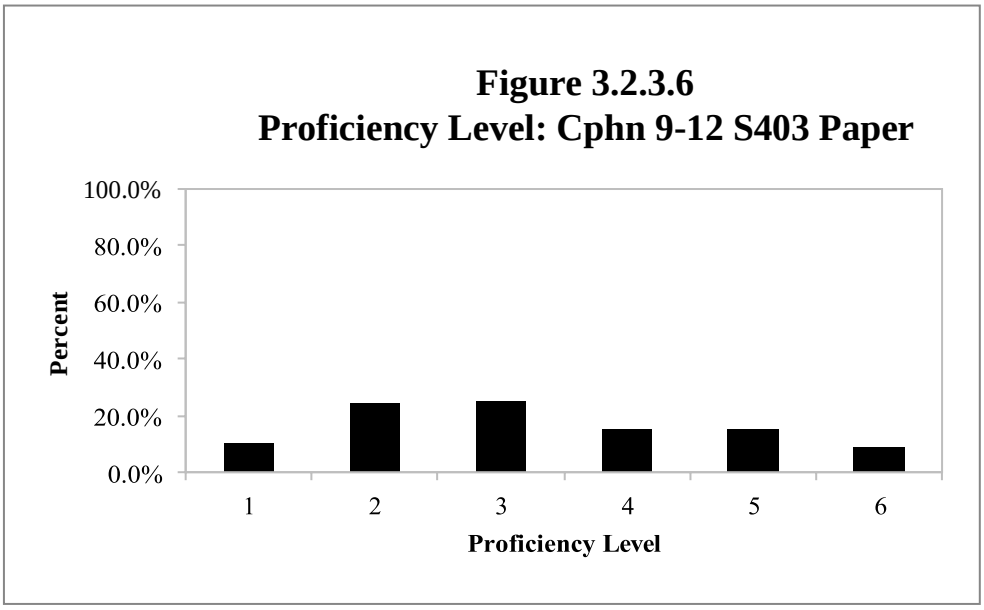
3.2.3.6 Grades 9–12

Table 3.2.3.6

Proficiency Level Distribution: Cphn 9-12 S403 Paper

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,494	10.3%	1,333	10.0%	1,364	11.5%	826	10.5%	5,017	10.6%
2	3,409	23.6%	2,976	22.4%	2,844	23.9%	2,343	29.9%	11,572	24.4%
3	3,469	24.0%	3,439	25.9%	2,909	24.5%	2,161	27.6%	11,978	25.2%
4	2,315	16.0%	2,216	16.7%	1,568	13.2%	1,149	14.7%	7,248	15.3%
5	2,475	17.1%	1,935	14.6%	1,966	16.5%	929	11.9%	7,305	15.4%
6	1,301	9.0%	1,370	10.3%	1,230	10.4%	424	5.4%	4,325	9.1%
Total	14,463	100.0%	13,269	100.0%	11,881	100.0%	7,832	100.0%	47,445	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

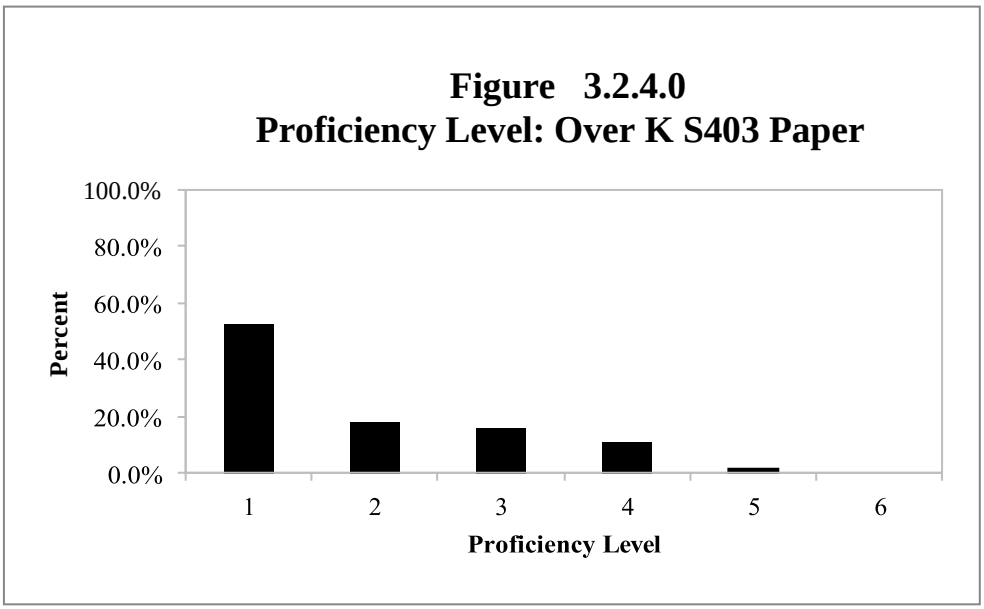
3.2.4 Overall

3.2.4.0 Kindergarten

Table 3.2.4.0

Proficiency Level Distribution: Over K S403 Paper

Level	Count	Percent
1	122,879	52.9%
2	42,312	18.2%
3	37,253	16.0%
4	25,935	11.2%
5	4,064	1.7%
6	0	0.0%
Total	232,443	100.0%



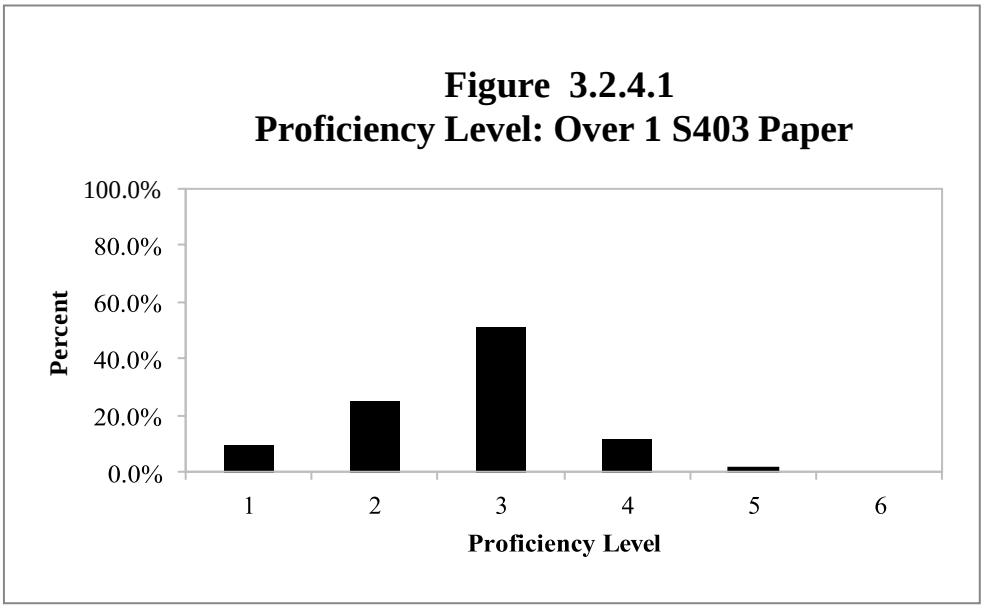
3.2.4.1 Grade 1

Table 3.2.4.1

Proficiency Level Distribution: Over 1 S403 Paper

Level	Count	Percent
1	3,625	9.8%
2	9,229	25.0%
3	18,877	51.2%
4	4,475	12.1%
5	631	1.7%
6	47	0.1%
Total	36,884	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

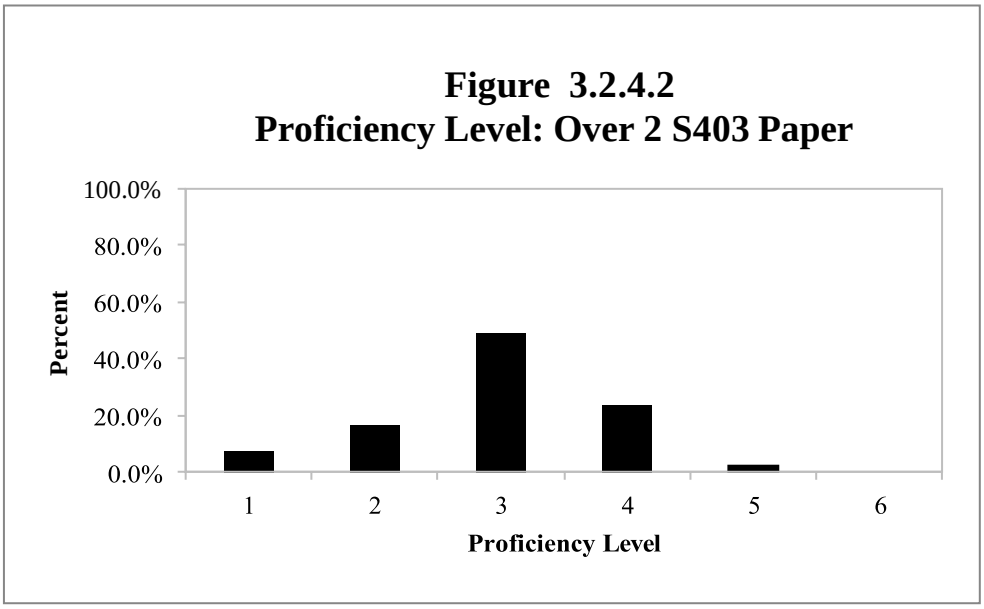
3.2.4.2 Grade 2

Table 3.2.4.2

Proficiency Level Distribution: Over 2 S403 Paper

Level	Count	Percent
1	2,897	7.6%
2	6,473	17.0%
3	18,745	49.2%
4	8,929	23.4%
5	1,026	2.7%
6	22	0.1%
Total	38,092	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

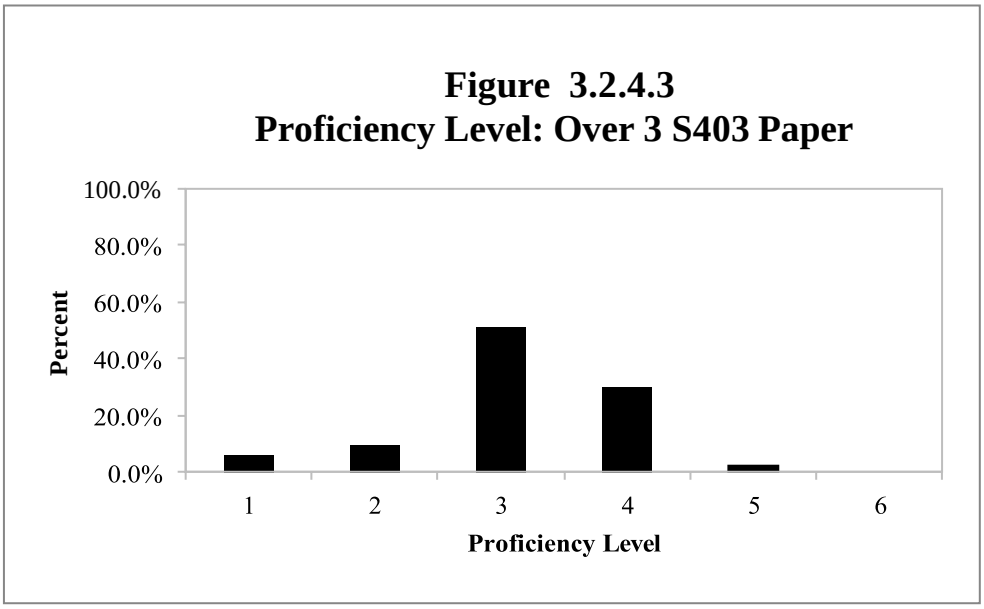
3.2.4.3 Grade 3

Table 3.2.4.3

Proficiency Level Distribution: Over 3 S403 Paper

Level	Count	Percent
1	1,972	5.9%
2	3,310	9.9%
3	16,959	50.9%
4	10,055	30.2%
5	969	2.9%
6	68	0.2%
Total	33,333	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

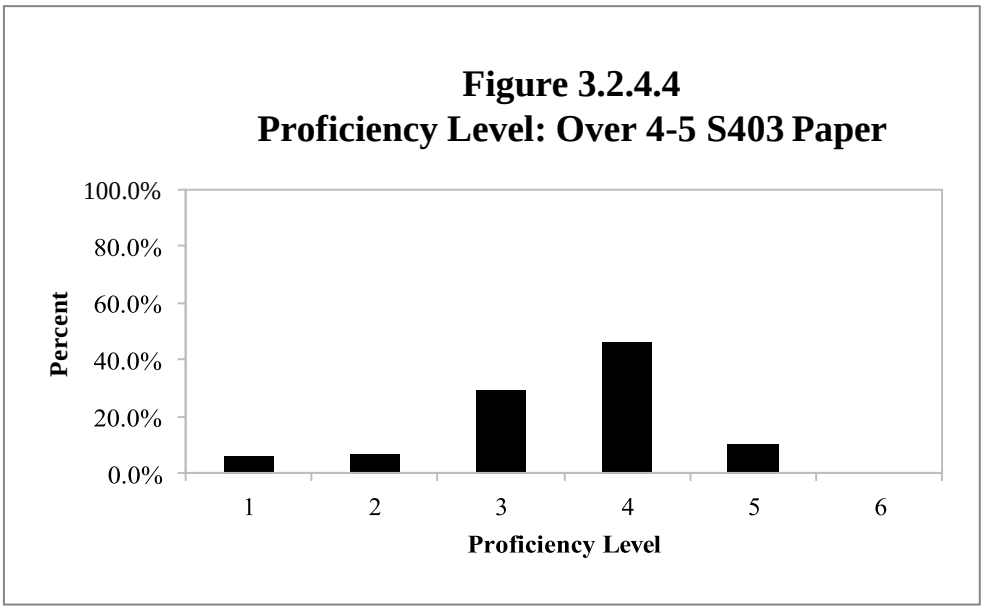
3.2.4.4 Grades 4–5

Table 3.2.4.4

Proficiency Level Distribution: Over 4-5 S403 Paper

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	1,616	6.0%	1,521	6.3%	3,137	6.2%
2	1,704	6.3%	1,645	6.8%	3,349	6.6%
3	8,414	31.3%	6,729	28.0%	15,143	29.7%
4	12,463	46.4%	10,950	45.5%	23,413	46.0%
5	2,428	9.0%	3,032	12.6%	5,460	10.7%
6	252	0.9%	195	0.8%	447	0.9%
Total	26,877	100.0%	24,072	100.0%	50,949	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

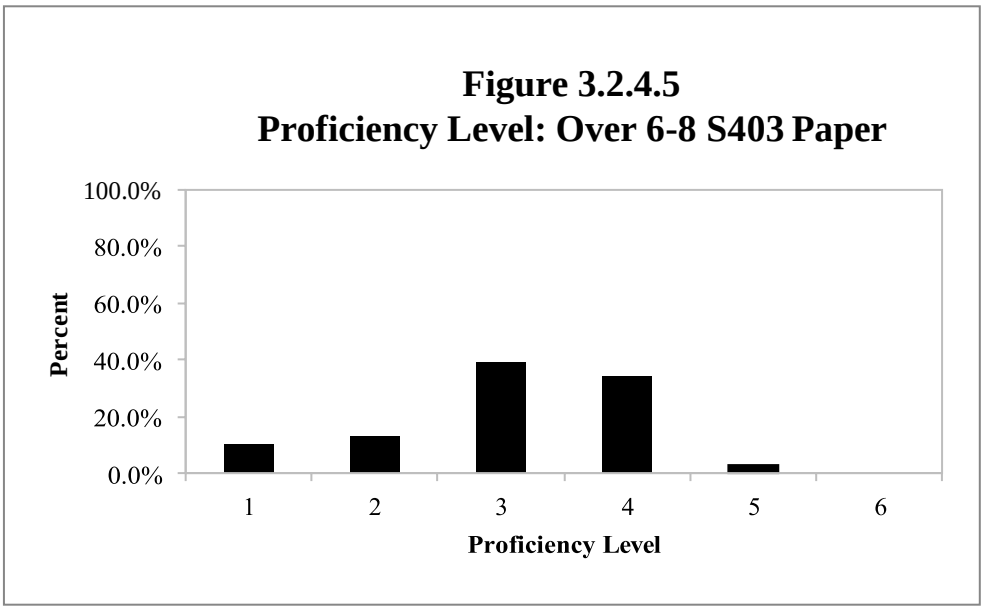
3.2.4.5 Grades 6–8

Table 3.2.4.5

Proficiency Level Distribution: Over 6-8 S403 Paper

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,420	8.2%	1,747	10.4%	1,869	12.1%	5,036	10.1%
2	2,185	12.6%	2,139	12.7%	2,179	14.2%	6,503	13.1%
3	7,226	41.5%	6,613	39.3%	5,633	36.6%	19,472	39.2%
4	6,070	34.9%	5,777	34.3%	5,188	33.7%	17,035	34.3%
5	479	2.8%	549	3.3%	508	3.3%	1,536	3.1%
6	13	0.1%	14	0.1%	8	0.1%	35	0.1%
Total	17,393	100.0%	16,839	100.0%	15,385	100.0%	49,617	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

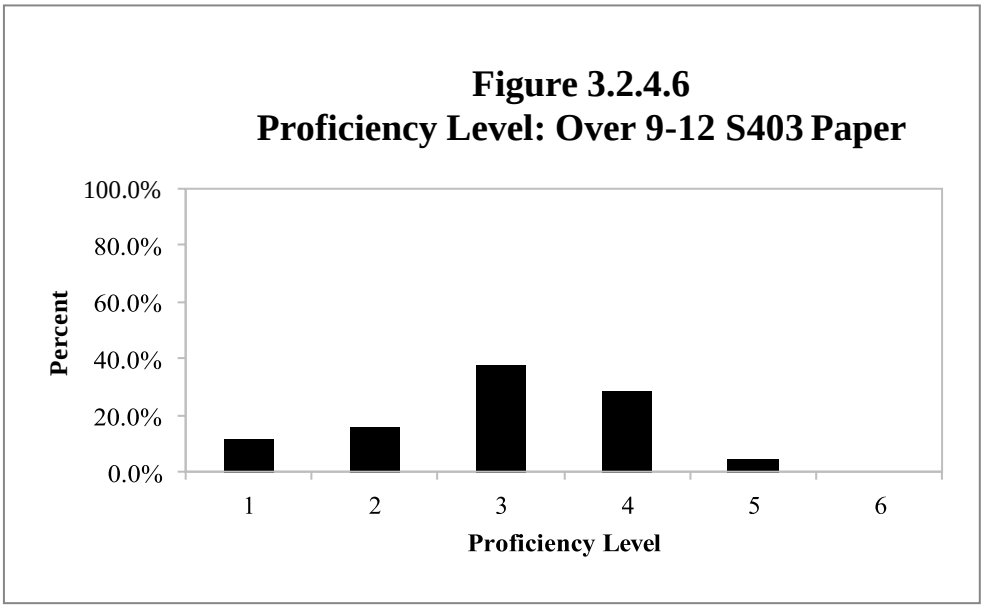
3.2.4.6 Grades 9–12

Table 3.2.4.6

Proficiency Level Distribution: Over 9-12 S403 Paper

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	1,850	12.9%	1,549	11.7%	1,434	12.2%	874	11.2%	5,707	12.1%
2	2,197	15.3%	2,048	15.5%	1,897	16.1%	1,510	19.4%	7,652	16.2%
3	4,998	34.8%	4,880	37.0%	4,456	37.8%	3,491	44.9%	17,825	37.8%
4	4,454	31.0%	3,988	30.2%	3,419	29.0%	1,759	22.6%	13,620	28.9%
5	854	5.9%	704	5.3%	583	4.9%	144	1.9%	2,285	4.8%
6	23	0.2%	17	0.1%	4	0.0%	0	0.0%	44	0.1%
Total	14,376	100.0%	13,186	100.0%	11,793	100.0%	7,778	100.0%	47,133	100.0%

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.



Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from the other states' results, they are excluded from this analysis.

3.3 Reliability

To estimate the score reliability of the composite scores, a stratified Cronbach's alpha coefficient (e.g., Rudner, 2001; Kamata, Turhan, & Darandari, 2003; Kane & Case, 2004) is computed, weighted by the contribution of each domain score into the composite. Specifically, the formula is

$$\alpha_c = 1 - \frac{\Sigma_{j=1}^{k} w_j^2 \sigma_j^2 (1-\rho_j)}{\sigma_c^2}$$

where

k = number of components j
w_j = weight of component j
o_j^2 = variance of component j
σ_c^2 = variance of composite

ρ_j = reliability coefficient of component j

Tables are provided below to express the stratified Cronbach's alpha for each of the composites. The first table for each composite provides stratified Cronbach's alpha for all test takers. The second table for each composite provides the same information for the population of female students and the population of male students. The third table provides information by ethnicity, for Hispanic and non-Hispanic test takers, and the fourth table provides information for the population of students who have an individualized education plan (IEP).

Each table is divided by grade-level cluster. Tables first include the input values used to compute Cronbach's alpha. The table lists the number of components for each composite and their weight. (See also the beginning of this chapter for an overview of how the composites are computed.)

For each grade-level cluster, a reliability coefficient is provided for each domain. To produce this coefficient, values for Cronbach's alpha for each of the tiers in the grade-level cluster (provided in Section 2.10) are weighted by the number of students who were administered the tier form, and a weighted average is expressed in the tables.

For each relevant domain component, the variance of the scale score is also provided. The variance of the composite scale score is also provided. The variances of domains and composites are computed for students who had valid results in all four domains.

Finally, the computed stratified Cronbach's alpha value for the composite is provided, by grade-level cluster.

The stratified Cronbach's alpha, presented in the tables in this section, was also used to produce the *Accuracy and Consistency* classification tables of the composites (Section 3.4).

3.3.1 Oral

Table 3.3.1.1

Reliability of Composite: Oral S403 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.50	5385.473	0.940
	Speaking	0.50	9879.079	0.899
	Oral		6641.816	0.950
1	Listening	0.50	1604.175	0.696
	Speaking	0.50	4064.308	0.894
	Oral		2078.799	0.890
2	Listening	0.50	1682.517	0.659
	Speaking	0.50	4249.091	0.911
	Oral		2206.207	0.892
3	Listening	0.50	1308.497	0.578
	Speaking	0.50	4326.814	0.909
	Oral		2023.187	0.883
4-5	Listening	0.50	1641.418	0.625
	Speaking	0.50	4898.727	0.905
	Oral		2473.967	0.891
6-8	Listening	0.50	2262.205	0.648
	Speaking	0.50	5533.852	0.908
	Oral		3105.455	0.895
9-12	Listening	0.50	2429.756	0.648
	Speaking	0.50	6156.818	0.921
	Oral		3421.574	0.902

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.1.2

Reliability of Composite: Oral S403 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.50	5088.414	0.938	5548.449	0.940
	Speaking	0.50	9841.417	0.901	9737.876	0.896
	Oral		6468.576	0.950	6660.987	0.950
1	Listening	0.50	1603.288	0.686	1595.786	0.701
	Speaking	0.50	4160.397	0.894	3951.775	0.894
	Oral		2111.285	0.888	2032.804	0.890
2	Listening	0.50	1640.018	0.644	1698.874	0.667
	Speaking	0.50	4179.839	0.911	4288.492	0.910
	Oral		2156.860	0.889	2228.673	0.893
3	Listening	0.50	1249.980	0.558	1353.001	0.594
	Speaking	0.50	4428.225	0.910	4232.483	0.908
	Oral		2031.067	0.883	2009.894	0.883
4-5	Listening	0.50	1591.661	0.617	1678.888	0.632
	Speaking	0.50	4945.436	0.904	4834.216	0.906
	Oral		2462.491	0.890	2471.305	0.891
6-8	Listening	0.50	2262.277	0.647	2258.361	0.649
	Speaking	0.50	5523.598	0.906	5529.525	0.910
	Oral		3114.424	0.894	3090.177	0.896
9-12	Listening	0.50	2317.155	0.637	2526.346	0.657
	Speaking	0.50	6108.518	0.921	6192.107	0.921
	Oral		3341.329	0.901	3487.622	0.903

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.1.3

Reliability of Composite: Oral S403 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.50	5477.737	0.940	4889.182	0.934
	Speaking	0.50	10081.665	0.901	9125.988	0.888
	Oral		6775.283	0.951	6039.095	0.944
1	Listening	0.50	1573.106	0.691	1684.430	0.708
	Speaking	0.50	4067.127	0.895	3908.678	0.886
	Oral		2057.563	0.889	2077.837	0.887
2	Listening	0.50	1685.124	0.651	1639.498	0.678
	Speaking	0.50	4268.611	0.911	4036.603	0.907
	Oral		2215.875	0.891	2096.496	0.892
3	Listening	0.50	1271.749	0.565	1413.671	0.617
	Speaking	0.50	4351.881	0.909	4094.886	0.906
	Oral		2012.339	0.882	1988.450	0.884
4-5	Listening	0.50	1647.782	0.617	1595.538	0.647
	Speaking	0.50	5013.342	0.905	4329.920	0.902
	Oral		2523.461	0.890	2219.124	0.889
6-8	Listening	0.50	2318.222	0.646	2023.853	0.655
	Speaking	0.50	5742.199	0.908	4587.212	0.902
	Oral		3224.858	0.896	2575.067	0.888
9-12	Listening	0.50	2425.389	0.644	2408.371	0.660
	Speaking	0.50	6349.845	0.923	5397.195	0.911
	Oral		3505.261	0.903	3074.664	0.895

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.1.4

Reliability of Composite: Oral S403 Paper by IEP status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.50	6375.405	0.949
	Speaking	0.50	8808.948	0.893
	Oral		6723.739	0.953
1	Listening	0.50	1709.996	0.715
	Speaking	0.50	3560.483	0.883
	Oral		1936.121	0.883
2	Listening	0.50	1701.517	0.700
	Speaking	0.50	3617.141	0.894
	Oral		1921.262	0.884
3	Listening	0.50	1083.813	0.548
	Speaking	0.50	3109.646	0.898
	Oral		1425.644	0.859
4-5	Listening	0.50	1270.935	0.617
	Speaking	0.50	3345.807	0.900
	Oral		1607.497	0.872
6-8	Listening	0.50	1560.830	0.615
	Speaking	0.50	3798.274	0.898
	Oral		1971.999	0.875
9-12	Listening	0.50	1712.923	0.612
	Speaking	0.50	4885.951	0.913
	Oral		2441.476	0.889

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.3.2 Literacy

Table 3.3.2.1

Reliability of Composite: Litr S403 Paper

Cluster	Component	Weight	Variance	Reliability
K	Reading	0.50	4552.769	0.950
	Writing	0.50	4448.735	0.925
	Literacy		3884.676	0.964
1	Reading	0.50	755.005	0.736
	Writing	0.50	1592.273	0.908
	Literacy		877.088	0.902
2	Reading	0.50	1035.865	0.826
	Writing	0.50	1523.765	0.933
	Literacy		1039.298	0.932
3	Reading	0.50	709.360	0.685
	Writing	0.50	1509.002	0.923
	Literacy		860.333	0.901
4-5	Reading	0.50	928.767	0.777
	Writing	0.50	1623.961	0.894
	Literacy		1048.224	0.909
6-8	Reading	0.50	867.827	0.784
	Writing	0.50	1670.437	0.903
	Literacy		1031.404	0.915
9-12	Reading	0.50	982.112	0.804
	Writing	0.50	1674.126	0.897
	Literacy		1096.009	0.917

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.2.2

Reliability of Composite: Litr S403 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Reading	0.50	4379.323	0.947	4695.109	0.953
	Writing	0.50	4297.693	0.924	4530.078	0.926
	Literacy		3733.978	0.962	3989.527	0.965
1	Reading	0.50	734.646	0.735	773.684	0.737
	Writing	0.50	1482.774	0.904	1654.071	0.910
	Literacy		835.462	0.899	903.699	0.903
2	Reading	0.50	1010.559	0.828	1052.597	0.824
	Writing	0.50	1392.523	0.926	1569.839	0.935
	Literacy		976.823	0.929	1065.927	0.933
3	Reading	0.50	673.856	0.672	736.714	0.694
	Writing	0.50	1466.083	0.915	1486.774	0.925
	Literacy		833.867	0.896	866.220	0.903
4-5	Reading	0.50	895.127	0.773	955.480	0.780
	Writing	0.50	1573.072	0.885	1619.310	0.896
	Literacy		1018.852	0.906	1057.705	0.910
6-8	Reading	0.50	845.679	0.776	880.973	0.787
	Writing	0.50	1605.270	0.891	1676.673	0.907
	Literacy		1000.262	0.909	1035.485	0.917
9-12	Reading	0.50	943.896	0.796	1004.811	0.808
	Writing	0.50	1662.448	0.892	1641.330	0.898
	Literacy		1078.336	0.914	1086.946	0.917

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.2.3

Reliability of Composite: Litr S403 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Reading	0.50	4197.795	0.946	4783.576	0.954
	Writing	0.50	4222.166	0.919	4473.469	0.929
	Literacy		3590.897	0.960	4015.183	0.967
1	Reading	0.50	714.886	0.719	852.908	0.775
	Writing	0.50	1565.913	0.906	1588.100	0.913
	Literacy		840.481	0.897	939.559	0.912
2	Reading	0.50	1012.840	0.822	1092.525	0.838
	Writing	0.50	1513.728	0.932	1519.089	0.935
	Literacy		1022.277	0.931	1071.683	0.936
3	Reading	0.50	684.256	0.665	777.854	0.731
	Writing	0.50	1509.799	0.922	1455.172	0.928
	Literacy		848.154	0.898	874.165	0.910
4-5	Reading	0.50	924.558	0.772	934.477	0.791
	Writing	0.50	1657.304	0.892	1477.042	0.899
	Literacy		1060.364	0.908	987.675	0.913
6-8	Reading	0.50	869.196	0.780	853.751	0.795
	Writing	0.50	1694.425	0.902	1546.990	0.905
	Literacy		1042.887	0.914	969.627	0.917
9-12	Reading	0.50	989.455	0.804	951.696	0.805
	Writing	0.50	1681.811	0.898	1621.918	0.894
	Literacy		1104.768	0.917	1052.246	0.915

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.2.4

Reliability of Composite: Litr S403 Paper by IEP status

Cluster	Component	Weight	Variance	Reliability
K	Reading	0.50	4649.744	0.959
	Writing	0.50	4264.734	0.924
	Literacy		3754.572	0.966
1	Reading	0.50	625.574	0.652
	Writing	0.50	1844.842	0.911
	Literacy		853.480	0.888
2	Reading	0.50	867.365	0.777
	Writing	0.50	1727.950	0.937
	Literacy		1010.924	0.925
3	Reading	0.50	561.185	0.550
	Writing	0.50	1402.444	0.934
	Literacy		716.404	0.880
4-5	Reading	0.50	652.340	0.714
	Writing	0.50	1257.703	0.906
	Literacy		725.670	0.895
6-8	Reading	0.50	596.558	0.709
	Writing	0.50	1327.040	0.908
	Literacy		719.487	0.897
9-12	Reading	0.50	673.158	0.741
	Writing	0.50	1364.912	0.907
	Literacy		776.691	0.903

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.3.3 Comprehension

Table 3.3.3.1

Reliability of Composite: Cphn S403 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.30	5385.473	0.940
	Reading	0.70	4552.769	0.950
	Comprehension		3797.434	0.963
1	Listening	0.30	1604.175	0.696
	Reading	0.70	755.005	0.736
	Comprehension		747.118	0.810
2	Listening	0.30	1682.517	0.659
	Reading	0.70	1035.865	0.826
	Comprehension		981.399	0.858
3	Listening	0.30	1308.497	0.578
	Reading	0.70	709.360	0.685
	Comprehension		719.711	0.779
4-5	Listening	0.30	1641.418	0.625
	Reading	0.70	928.767	0.777
	Comprehension		969.129	0.838
6-8	Listening	0.30	2262.205	0.648
	Reading	0.70	867.827	0.784
	Comprehension		1048.086	0.844
9-12	Listening	0.30	2429.756	0.648
	Reading	0.70	982.112	0.804
	Comprehension		1168.688	0.853

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.3.2

Reliability of Composite: Cphn S403 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.30	5088.414	0.938	5548.449	0.940
	Reading	0.70	4379.323	0.947	4695.109	0.953
	Comprehension		3650.493	0.961	3897.822	0.964
1	Listening	0.30	1603.288	0.686	1595.786	0.701
	Reading	0.70	734.646	0.735	773.684	0.737
	Comprehension		737.732	0.809	753.962	0.811
2	Listening	0.30	1640.018	0.644	1698.874	0.667
	Reading	0.70	1010.559	0.828	1052.597	0.824
	Comprehension		954.971	0.856	995.530	0.858
3	Listening	0.30	1249.980	0.558	1353.001	0.594
	Reading	0.70	673.856	0.672	736.714	0.694
	Comprehension		682.029	0.768	748.310	0.786
4-5	Listening	0.30	1591.661	0.617	1678.888	0.632
	Reading	0.70	895.127	0.773	955.480	0.780
	Comprehension		935.893	0.835	994.774	0.840
6-8	Listening	0.30	2262.277	0.647	2258.361	0.649
	Reading	0.70	845.679	0.776	880.973	0.787
	Comprehension		1038.853	0.842	1051.973	0.845
9-12	Listening	0.30	2317.155	0.637	2526.346	0.657
	Reading	0.70	943.896	0.796	1004.811	0.808
	Comprehension		1119.970	0.848	1205.570	0.857

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.3.3

Reliability of Composite: Cphn S403 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.30	5477.737	0.940	4889.182	0.934
	Reading	0.70	4197.795	0.946	4783.576	0.954
	Comprehension		3567.123	0.961	3852.103	0.964
1	Listening	0.30	1573.106	0.691	1684.430	0.708
	Reading	0.70	714.886	0.719	852.908	0.775
	Comprehension		709.362	0.800	844.377	0.836
2	Listening	0.30	1685.124	0.651	1639.498	0.678
	Reading	0.70	1012.840	0.822	1092.525	0.838
	Comprehension		962.079	0.853	1025.580	0.869
3	Listening	0.30	1271.749	0.565	1413.671	0.617
	Reading	0.70	684.256	0.665	777.854	0.731
	Comprehension		691.614	0.766	797.647	0.810
4-5	Listening	0.30	1647.782	0.617	1595.538	0.647
	Reading	0.70	924.558	0.772	934.477	0.791
	Comprehension		968.007	0.835	960.092	0.848
6-8	Listening	0.30	2318.222	0.646	2023.853	0.655
	Reading	0.70	869.196	0.780	853.751	0.795
	Comprehension		1060.982	0.842	985.073	0.849
9-12	Listening	0.30	2425.389	0.644	2408.371	0.660
	Reading	0.70	989.455	0.804	951.696	0.805
	Comprehension		1173.833	0.853	1139.290	0.856

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.3.4

Reliability of Composite: Cphn S403 Paper by IEP status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.30	6375.405	0.949
	Reading	0.70	4649.744	0.959
	Comprehension		3909.026	0.969
1	Listening	0.30	1709.996	0.715
	Reading	0.70	625.574	0.652
	Comprehension		650.702	0.769
2	Listening	0.30	1701.517	0.700
	Reading	0.70	867.365	0.777
	Comprehension		843.999	0.833
3	Listening	0.30	1083.813	0.548
	Reading	0.70	561.185	0.550
	Comprehension		538.269	0.688
4-5	Listening	0.30	1270.935	0.617
	Reading	0.70	652.340	0.714
	Comprehension		657.530	0.795
6-8	Listening	0.30	1560.830	0.615
	Reading	0.70	596.558	0.709
	Comprehension		683.801	0.797
9-12	Listening	0.30	1712.923	0.612
	Reading	0.70	673.158	0.741
	Comprehension		771.064	0.811

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.3.4 Overall

Table 3.3.4.1

Reliability of Composite: Over S403 Paper

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.15	5385.473	0.940
	Reading	0.35	4552.769	0.950
	Speaking	0.15	9879.079	0.899
	Writing	0.35	4448.735	0.925
	Overall Composite		3760.516	0.974
1	Listening	0.15	1604.175	0.696
	Reading	0.35	755.005	0.736
	Speaking	0.15	4064.308	0.894
	Writing	0.35	1592.273	0.908
	Overall Composite		966.370	0.935
2	Listening	0.15	1682.517	0.659
	Reading	0.35	1035.865	0.826
	Speaking	0.15	4249.091	0.911
	Writing	0.35	1523.765	0.933
	Overall Composite		1128.691	0.950
3	Listening	0.15	1308.497	0.578
	Reading	0.35	709.360	0.685
	Speaking	0.15	4326.814	0.909
	Writing	0.35	1509.002	0.923
	Overall Composite		987.257	0.936
4-5	Listening	0.15	1641.418	0.625
	Reading	0.35	928.767	0.777
	Speaking	0.15	4898.727	0.905
	Writing	0.35	1623.961	0.894
	Overall Composite		1246.619	0.943
6-8	Listening	0.15	2262.205	0.648
	Reading	0.35	867.827	0.784
	Speaking	0.15	5533.852	0.908
	Writing	0.35	1670.437	0.903
	Overall Composite		1376.922	0.948
9-12	Listening	0.15	2429.756	0.648
	Reading	0.35	982.112	0.804
	Speaking	0.15	6156.818	0.921
	Writing	0.35	1674.126	0.897
	Overall Composite		1481.605	0.949

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.4.2

Reliability of Composite: Over S403 Paper by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
K	Listening	0.15	5088.414	0.938	5548.449	0.940
	Reading	0.35	4379.323	0.947	4695.109	0.953
	Speaking	0.15	9841.417	0.901	9737.876	0.896
	Writing	0.35	4297.693	0.924	4530.078	0.926
	Overall Composite		3629.599	0.973	3824.205	0.974
1	Listening	0.15	1603.288	0.686	1595.786	0.701
	Reading	0.35	734.646	0.735	773.684	0.737
	Speaking	0.15	4160.397	0.894	3951.775	0.894
	Writing	0.35	1482.774	0.904	1654.071	0.910
	Overall Composite		945.437	0.934	972.848	0.935
2	Listening	0.15	1640.018	0.644	1698.874	0.667
	Reading	0.35	1010.559	0.828	1052.597	0.824
	Speaking	0.15	4179.839	0.911	4288.492	0.910
	Writing	0.35	1392.523	0.926	1569.839	0.935
	Overall Composite		1074.636	0.949	1151.563	0.951
3	Listening	0.15	1249.980	0.558	1353.001	0.594
	Reading	0.35	673.856	0.672	736.714	0.694
	Speaking	0.15	4428.225	0.910	4232.483	0.908
	Writing	0.35	1466.083	0.915	1486.774	0.925
	Overall Composite		975.803	0.935	986.907	0.937
4-5	Listening	0.15	1591.661	0.617	1678.888	0.632
	Reading	0.35	895.127	0.773	955.480	0.780
	Speaking	0.15	4945.436	0.904	4834.216	0.906
	Writing	0.35	1573.072	0.885	1619.310	0.896
	Overall Composite		1225.924	0.942	1251.535	0.944
6-8	Listening	0.15	2262.277	0.647	2258.361	0.649
	Reading	0.35	845.679	0.776	880.973	0.787
	Speaking	0.15	5523.598	0.906	5529.525	0.910
	Writing	0.35	1605.270	0.891	1676.673	0.907
	Overall Composite		1366.204	0.946	1372.952	0.948
9-12	Listening	0.15	2317.155	0.637	2526.346	0.657
	Reading	0.35	943.896	0.796	1004.811	0.808
	Speaking	0.15	6108.518	0.921	6192.107	0.921
	Writing	0.35	1662.448	0.892	1641.330	0.898
	Overall Composite		1463.766	0.948	1484.648	0.950

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.4.3

Reliability of Composite: Over S403 Paper by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
K	Listening	0.15	5477.737	0.940	4889.182	0.934
	Reading	0.35	4197.795	0.946	4783.576	0.954
	Speaking	0.15	10081.665	0.901	9125.988	0.888
	Writing	0.35	4222.166	0.919	4473.469	0.929
	Overall Composite		3571.604	0.972	3734.846	0.974
1	Listening	0.15	1573.106	0.691	1684.430	0.708
	Reading	0.35	714.886	0.719	852.908	0.775
	Speaking	0.15	4067.127	0.895	3908.678	0.886
	Writing	0.35	1565.913	0.906	1588.100	0.913
	Overall Composite		930.291	0.932	1024.766	0.940
2	Listening	0.15	1685.124	0.651	1639.498	0.678
	Reading	0.35	1012.840	0.822	1092.525	0.838
	Speaking	0.15	4268.611	0.911	4036.603	0.907
	Writing	0.35	1513.728	0.932	1519.089	0.935
	Overall Composite		1115.533	0.949	1136.085	0.952
3	Listening	0.15	1271.749	0.565	1413.671	0.617
	Reading	0.35	684.256	0.665	777.854	0.731
	Speaking	0.15	4351.881	0.909	4094.886	0.906
	Writing	0.35	1509.799	0.922	1455.172	0.928
	Overall Composite		976.263	0.935	987.883	0.940
4-5	Listening	0.15	1647.782	0.617	1595.538	0.647
	Reading	0.35	924.558	0.772	934.477	0.791
	Speaking	0.15	5013.342	0.905	4329.920	0.902
	Writing	0.35	1657.304	0.892	1477.042	0.899
	Overall Composite		1268.328	0.943	1138.525	0.943
6-8	Listening	0.15	2318.222	0.646	2023.853	0.655
	Reading	0.35	869.196	0.780	853.751	0.795
	Speaking	0.15	5742.199	0.908	4587.212	0.902
	Writing	0.35	1694.425	0.902	1546.990	0.905
	Overall Composite		1413.913	0.948	1207.735	0.946
9-12	Listening	0.15	2425.389	0.644	2408.371	0.660
	Reading	0.35	989.455	0.804	951.696	0.805
	Speaking	0.15	6349.845	0.923	5397.195	0.911
	Writing	0.35	1681.811	0.898	1621.918	0.894
	Overall Composite		1509.516	0.950	1363.138	0.946

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.3.4.4

Reliability of Composite: Over S403 Paper by IEP status

Cluster	Component	Weight	Variance	Reliability
K	Listening	0.15	6375.405	0.949
	Reading	0.35	4649.744	0.959
	Speaking	0.15	8808.948	0.893
	Writing	0.35	4264.734	0.924
	Overall Composite		3659.532	0.975
1	Listening	0.15	1709.996	0.715
	Reading	0.35	625.574	0.652
	Speaking	0.15	3560.483	0.883
	Writing	0.35	1844.842	0.911
	Overall Composite		907.169	0.926
2	Listening	0.15	1701.517	0.700
	Reading	0.35	867.365	0.777
	Speaking	0.15	3617.141	0.894
	Writing	0.35	1727.950	0.937
	Overall Composite		1024.457	0.944
3	Listening	0.15	1083.813	0.548
	Reading	0.35	561.185	0.550
	Speaking	0.15	3109.646	0.898
	Writing	0.35	1402.444	0.934
	Overall Composite		710.364	0.915
4-5	Listening	0.15	1270.935	0.617
	Reading	0.35	652.340	0.714
	Speaking	0.15	3345.807	0.900
	Writing	0.35	1257.703	0.906
	Overall Composite		777.087	0.928
6-8	Listening	0.15	1560.830	0.615
	Reading	0.35	596.558	0.709
	Speaking	0.15	3798.274	0.898
	Writing	0.35	1327.040	0.908
	Overall Composite		862.663	0.932
9-12	Listening	0.15	1712.923	0.612
	Reading	0.35	673.158	0.741
	Speaking	0.15	4885.951	0.913
	Writing	0.35	1364.912	0.907
	Overall Composite		996.635	0.938

Note: Student scale scores from the state of Massachusetts and Colorado are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.4 Accuracy and Consistency of Composites

Tables below present three sections of information related to the accuracy and consistency of placement into the WIDA language proficiency levels for each composite score. The first section provides overall indices related to the accuracy and consistency of classification, as well as Cohen's kappa. The second section shows accuracy and consistency information conditional per proficiency level. The third section provides indices of classification accuracy, including the false-positives and false-negatives, and consistency at the cut points. These indices are perhaps the most important of all when using any of these as an absolute cut point (e.g., determining which students have reached PL 5). Note that the consistency is generally higher at the cut points than for the proficiency levels. For practical purposes, the primary score used for such decisions is the overall composite score.

As noted above in Section 2.13, there may be cases where the number of test takers placed into the proficiency level is fewer than 200 and accuracy and consistency of classification conditional on that level either cannot be computed or the software produces estimates that are out of bounds. In these cases, 'N/A' has been placed in the table. In addition, there may be cases where due to the small percentage of test takers placed into the proficiency level and the range of observed scale scores, accuracy of classification conditional on that level cannot be estimated by the software program that is used. In such cases, a hyphen (-) has been placed in the table.

3.4.1 Oral

Table 3.4.1.0

Accuracy and Consistency of Classification Indices: Oral (Grade K) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.702	0.614	0.528
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.888	0.835
	2	0.624	0.507
	3	0.532	0.413
	4	0.430	0.321
	5	0.627	0.509
	6	0.845	0.765

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.946	0.028	0.026	0.924
	2/3	0.937	0.027	0.036	0.913
	3/4	0.941	0.029	0.030	0.914
	4/5	0.934	0.038	0.028	0.907
	5/6	0.931	0.036	0.032	0.903

Table 3.4.1.1

Accuracy and Consistency of Classification Indices: Oral (Grade 1) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.659	0.547	0.416
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.827	0.690
	2	0.626	0.491
	3	0.716	0.614
	4	0.625	0.516
	5	0.610	0.524
	6	0.670	0.382

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.975	0.009	0.016	0.964
	2/3	0.936	0.031	0.033	0.907
	3/4	0.888	0.051	0.061	0.845
	4/5	0.903	0.048	0.049	0.862
	5/6	0.954	0.044	0.002	0.944

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.2

Accuracy and Consistency of Classification Indices: Oral (Grade 2) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.646	0.538	0.396
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.869	0.763
	2	0.595	0.458
	3	0.727	0.616
	4	0.672	0.559
	5	0.500	0.426
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.979	0.007	0.015	0.970
	2/3	0.945	0.030	0.025	0.919
	3/4	0.887	0.047	0.066	0.844
	4/5	0.887	0.044	0.069	0.835
	5/6	0.946	0.054	0.000	0.941

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.3

Accuracy and Consistency of Classification Indices: Oral (Grade 3) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.663	0.553	0.407
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.878	0.770
	2	0.584	0.446
	3	0.736	0.627
	4	0.682	0.579
	5	0.537	0.449
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.980	0.006	0.015	0.972
	2/3	0.947	0.029	0.024	0.921
	3/4	0.885	0.048	0.067	0.841
	4/5	0.893	0.049	0.058	0.845
	5/6	0.956	0.044	0.000	0.949

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.4

Accuracy and Consistency of Classification Indices: Oral (Grade 4) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.620	0.519		0.376	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.891		0.797	
	2	0.563		0.426	
	3	0.656		0.526	
	4	0.713		0.595	
	5	0.510		0.456	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.981	0.005	0.014	0.974
	2/3	0.962	0.019	0.019	0.943
	3/4	0.911	0.042	0.047	0.876
	4/5	0.869	0.043	0.088	0.817
	5/6	0.893	0.107	0.000	0.875

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.5

Accuracy and Consistency of Classification Indices: Oral (Grade 5) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.624	0.518		0.370	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.890		0.798	
	2	0.541		0.406	
	3	0.622		0.490	
	4	0.722		0.604	
	5	0.520		0.462	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.979	0.006	0.015	0.970
	2/3	0.961	0.020	0.019	0.942
	3/4	0.912	0.043	0.045	0.877
	4/5	0.863	0.045	0.091	0.808
	5/6	0.905	0.095	0.000	0.882

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.6

Accuracy and Consistency of Classification Indices: Oral (Grade 6) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.621	0.516	0.385	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.862	0.758	
	2	0.587	0.454	
	3	0.644	0.522	
	4	0.687	0.571	
	5	0.499	0.437	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.972	0.010	0.018	0.960
	2/3	0.947	0.025	0.028	0.923
	3/4	0.904	0.045	0.051	0.866
	4/5	0.880	0.048	0.073	0.828
	5/6	0.913	0.087	0.000	0.896

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.7

Accuracy and Consistency of Classification Indices: Oral (Grade 7) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.620	0.514	0.388	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.858	0.756	
	2	0.559	0.431	
	3	0.612	0.492	
	4	0.684	0.575	
	5	0.497	0.428	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.964	0.013	0.023	0.949
	2/3	0.938	0.028	0.033	0.911
	3/4	0.899	0.050	0.052	0.858
	4/5	0.889	0.048	0.062	0.842
	5/6	0.924	0.076	0.000	0.908

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.8

Accuracy and Consistency of Classification Indices: Oral (Grade 8) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.613	0.507	0.385	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.854	0.754	
	2	0.561	0.435	
	3	0.596	0.478	
	4	0.670	0.558	
	5	0.481	0.411	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**		
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.958	0.016	0.026	0.940
	2/3	0.933	0.030	0.037	0.904
	3/4	0.900	0.050	0.051	0.860
	4/5	0.890	0.050	0.060	0.843
	5/6	0.925	0.075	0.000	0.907

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.9

Accuracy and Consistency of Classification Indices: Oral (Grade 9) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.650	0.542	0.427	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.871	0.790	
	2	0.545	0.425	
	3	0.640	0.526	
	4	0.659	0.551	
	5	0.510	0.416	
	6	-	-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**		
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.944	0.023	0.034	0.921
	2/3	0.922	0.039	0.039	0.889
	3/4	0.901	0.050	0.050	0.861
	4/5	0.915	0.045	0.040	0.879
	5/6	0.961	0.039	0.000	0.946

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.10

Accuracy and Consistency of Classification Indices: Oral (Grade 10) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.656	0.549	0.433
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.871	0.785
	2	0.531	0.412
	3	0.666	0.559
	4	0.671	0.563
	5	0.476	0.379
	6	0.664	0.382

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.944	0.021	0.034	0.921
	2/3	0.919	0.040	0.041	0.885
	3/4	0.898	0.053	0.050	0.857
	4/5	0.928	0.040	0.032	0.895
	5/6	0.960	0.038	0.003	0.949

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.11

Accuracy and Consistency of Classification Indices: Oral (Grade 11) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.665	0.557	0.439
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.862	0.773
	2	0.525	0.407
	3	0.674	0.570
	4	0.677	0.570
	5	0.513	0.404
	6	0.680	0.382

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.942	0.023	0.035	0.918
	2/3	0.917	0.040	0.042	0.883
	3/4	0.898	0.054	0.048	0.858
	4/5	0.932	0.039	0.029	0.902
	5/6	0.968	0.030	0.002	0.959

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.1.12

Accuracy and Consistency of Classification Indices: Oral (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.701	0.596	0.472
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.851	0.754
	2	0.545	0.426
	3	0.727	0.638
	4	0.725	0.620
	5	0.503	0.378
	6	0.804	0.518

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.944	0.022	0.034	0.920
	2/3	0.915	0.041	0.044	0.880
	3/4	0.902	0.053	0.045	0.862
	4/5	0.956	0.027	0.017	0.934
	5/6	0.980	0.019	0.002	0.975

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.4.2 Literacy

Table 3.4.2.0

Accuracy and Consistency of Classification Indices: Litr (Grade K) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.852	0.816	0.615
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.964	0.950
	2	0.646	0.518
	3	0.559	0.512
	4	-	-
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.949	0.025	0.026	0.928
	2/3	0.954	0.019	0.027	0.933
	3/4	0.949	0.051	0.000	0.946
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.1

Accuracy and Consistency of Classification Indices: Litr (Grade 1) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.790	0.706	0.569
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.827	0.731
	2	0.781	0.701
	3	0.802	0.735
	4	0.624	0.482
	5	0.889	0.538
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.933	0.032	0.034	0.904
	2/3	0.897	0.049	0.054	0.856
	3/4	0.967	0.022	0.010	0.952
	4/5	0.993	0.007	0.000	0.993
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.2

Accuracy and Consistency of Classification Indices: Litr (Grade 2) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.810	0.736		0.607	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.856		0.774	
	2	0.777		0.684	
	3	0.846		0.796	
	4	0.707		0.603	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.964	0.017	0.019	0.948
	2/3	0.926	0.034	0.040	0.896
	3/4	0.931	0.037	0.032	0.903
	4/5	0.988	0.012	0.000	0.988
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.3

Accuracy and Consistency of Classification Indices: Litr (Grade 3) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.782	0.701		0.504	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.880		0.800	
	2	0.684		0.558	
	3	0.860		0.806	
	4	0.612		0.511	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.978	0.008	0.014	0.969
	2/3	0.942	0.029	0.028	0.917
	3/4	0.876	0.053	0.071	0.827
	4/5	0.985	0.015	0.000	0.985
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.4

Accuracy and Consistency of Classification Indices: Litr (Grade 4) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.777	0.696		0.538	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.892		0.813	
	2	0.594		0.457	
	3	0.851		0.777	
	4	0.727		0.678	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.980	0.007	0.013	0.972
	2/3	0.959	0.021	0.020	0.940
	3/4	0.887	0.041	0.071	0.842
	4/5	0.951	0.049	0.000	0.938
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.5

Accuracy and Consistency of Classification Indices: Litr (Grade 5) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.750	0.664		0.509	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.880		0.793	
	2	0.628		0.494	
	3	0.828		0.746	
	4	0.701		0.664	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.980	0.007	0.013	0.972
	2/3	0.958	0.021	0.022	0.939
	3/4	0.897	0.040	0.063	0.856
	4/5	0.915	0.085	0.000	0.894
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.6

Accuracy and Consistency of Classification Indices: Litr (Grade 6) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.815	0.740	0.595
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.874	0.785
	2	0.714	0.604
	3	0.851	0.804
	4	0.788	0.682
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.970	0.011	0.020	0.958
	2/3	0.932	0.033	0.035	0.904
	3/4	0.914	0.046	0.040	0.878
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.7

Accuracy and Consistency of Classification Indices: Litr (Grade 7) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.801	0.720	0.577
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.883	0.802
	2	0.720	0.611
	3	0.839	0.782
	4	0.742	0.638
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.968	0.012	0.020	0.956
	2/3	0.933	0.033	0.034	0.904
	3/4	0.900	0.047	0.052	0.860
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.8

Accuracy and Consistency of Classification Indices: Litr (Grade 8) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.796	0.713	0.583
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.887	0.811
	2	0.687	0.575
	3	0.826	0.764
	4	0.776	0.679
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			
		Accuracy	False Positives	False Negatives	Consistency
	1/2	0.962	0.014	0.025	0.946
	2/3	0.929	0.036	0.035	0.899
	3/4	0.906	0.046	0.048	0.867
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.9

Accuracy and Consistency of Classification Indices: Litr (Grade 9) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.770	0.681	0.557
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.872	0.780
	2	0.714	0.606
	3	0.804	0.734
	4	0.740	0.666
	5	0.647	0.412
	6	N/A	N/A

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			
		Accuracy	False Positives	False Negatives	Consistency
	1/2	0.967	0.011	0.021	0.954
	2/3	0.934	0.031	0.035	0.906
	3/4	0.910	0.045	0.045	0.874
	4/5	0.959	0.032	0.008	0.945
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.10

Accuracy and Consistency of Classification Indices: Litr (Grade 10) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.780	0.692	0.566
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.865	0.766
	2	0.712	0.604
	3	0.811	0.745
	4	0.764	0.685
	5	0.680	0.455
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.969	0.011	0.020	0.957
	2/3	0.934	0.032	0.035	0.906
	3/4	0.912	0.047	0.042	0.876
	4/5	0.966	0.025	0.009	0.952
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.11

Accuracy and Consistency of Classification Indices: Litr (Grade 11) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.772	0.683	0.559
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.862	0.765
	2	0.733	0.628
	3	0.797	0.727
	4	0.746	0.674
	5	0.629	0.383
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.968	0.012	0.020	0.954
	2/3	0.932	0.031	0.037	0.903
	3/4	0.911	0.046	0.043	0.875
	4/5	0.963	0.031	0.007	0.949
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.2.12

Accuracy and Consistency of Classification Indices: Litr (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.811	0.734	0.607
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.854	0.753
	2	0.737	0.639
	3	0.834	0.781
	4	0.826	0.729
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.963	0.014	0.023	0.948
	2/3	0.922	0.038	0.040	0.889
	3/4	0.927	0.041	0.032	0.896
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.4.3 Comprehension

Table 3.4.3.0

Accuracy and Consistency of Classification Indices: Cphn (Grade K) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.815	0.761	0.592
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.960	0.945
	2	0.477	0.361
	3	0.561	0.443
	4	0.368	0.277
	5	0.743	0.652
	6	0.744	0.504

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.949	0.025	0.026	0.928
	2/3	0.954	0.025	0.021	0.934
	3/4	0.956	0.028	0.017	0.939
	4/5	0.958	0.025	0.017	0.942
	5/6	0.982	0.014	0.004	0.974

Table 3.4.3.1

Accuracy and Consistency of Classification Indices: Cphn (Grade 1) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.555	0.444	0.302
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.688	0.414
	2	0.510	0.400
	3	0.596	0.498
	4	0.352	0.269
	5	0.573	0.447
	6	0.783	0.614

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.949	0.008	0.042	0.930
	2/3	0.883	0.057	0.061	0.833
	3/4	0.856	0.078	0.066	0.802
	4/5	0.881	0.067	0.051	0.833
	5/6	0.944	0.036	0.020	0.918

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.2

Accuracy and Consistency of Classification Indices: Cphn (Grade 2) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.587	0.477	0.357
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.784	0.600
	2	0.612	0.484
	3	0.584	0.474
	4	0.371	0.285
	5	0.590	0.475
	6	0.781	0.631

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.965	0.010	0.024	0.950
	2/3	0.909	0.041	0.050	0.871
	3/4	0.875	0.060	0.065	0.826
	4/5	0.881	0.067	0.052	0.833
	5/6	0.927	0.045	0.028	0.894

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.3

Accuracy and Consistency of Classification Indices: Cphn (Grade 3) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.533	0.420	0.266
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.847	0.646
	2	0.542	0.388
	3	0.545	0.423
	4	0.416	0.333
	5	0.541	0.449
	6	0.642	0.445

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.978	0.003	0.018	0.972
	2/3	0.938	0.025	0.037	0.905
	3/4	0.848	0.077	0.075	0.790
	4/5	0.823	0.081	0.096	0.763
	5/6	0.900	0.061	0.039	0.852

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.4

Accuracy and Consistency of Classification Indices: Cphn (Grade 4) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)		
	0.569	0.456	0.318		
Conditional on Level	**Level**	**Accuracy**	**Consistency**		
	1	0.892	0.766		
	2	0.543	0.404		
	3	0.596	0.466		
	4	0.420	0.329		
	5	0.576	0.481		
	6	0.651	0.494		
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.977	0.004	0.019	0.970
	2/3	0.947	0.026	0.027	0.919
	3/4	0.876	0.060	0.065	0.827
	4/5	0.849	0.070	0.081	0.794
	5/6	0.889	0.059	0.053	0.841

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.5

Accuracy and Consistency of Classification Indices: Cphn (Grade 5) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)		
	0.560	0.451	0.320		
Conditional on Level	**Level**	**Accuracy**	**Consistency**		
	1	0.872	0.735		
	2	0.523	0.387		
	3	0.554	0.431		
	4	0.410	0.321		
	5	0.534	0.438		
	6	0.710	0.559		
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.971	0.006	0.023	0.962
	2/3	0.938	0.030	0.032	0.906
	3/4	0.872	0.064	0.065	0.823
	4/5	0.852	0.068	0.080	0.798
	5/6	0.888	0.063	0.049	0.840

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.6

Accuracy and Consistency of Classification Indices: Cphn (Grade 6) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)	
	0.591	0.479	0.345	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.838	0.683	
	2	0.646	0.524	
	3	0.613	0.512	
	4	0.455	0.359	
	5	0.532	0.416	
	6	0.724	0.489	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**		
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.964	0.009	0.027	0.952
	2/3	0.904	0.047	0.049	0.862
	3/4	0.861	0.075	0.065	0.810
	4/5	0.887	0.058	0.054	0.841
	5/6	0.950	0.038	0.012	0.929

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.7

Accuracy and Consistency of Classification Indices: Cphn (Grade 7) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)	
	0.571	0.460	0.334	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	1	0.841	0.698	
	2	0.606	0.486	
	3	0.584	0.479	
	4	0.416	0.325	
	5	0.506	0.393	
	6	0.705	0.488	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**		
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.953	0.012	0.035	0.935
	2/3	0.898	0.051	0.051	0.854
	3/4	0.863	0.071	0.066	0.813
	4/5	0.885	0.062	0.053	0.838
	5/6	0.941	0.042	0.018	0.914

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.8

Accuracy and Consistency of Classification Indices: Cphn (Grade 8) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.567	0.456	0.333
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.828	0.689
	2	0.597	0.480
	3	0.551	0.443
	4	0.374	0.290
	5	0.560	0.444
	6	0.658	0.449

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.944	0.017	0.039	0.922
	2/3	0.895	0.052	0.052	0.851
	3/4	0.870	0.068	0.062	0.820
	4/5	0.881	0.068	0.051	0.836
	5/6	0.939	0.039	0.022	0.910

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.9

Accuracy and Consistency of Classification Indices: Cphn (Grade 9) S403 Paper

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.580	0.471	0.349
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.811	0.659
	2	0.659	0.546
	3	0.535	0.430
	4	0.397	0.309
	5	0.534	0.418
	6	0.719	0.526

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.949	0.016	0.035	0.929
	2/3	0.893	0.050	0.057	0.848
	3/4	0.873	0.070	0.057	0.825
	4/5	0.888	0.064	0.048	0.845
	5/6	0.942	0.037	0.021	0.915

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.10

Accuracy and Consistency of Classification Indices: Cphn (Grade 10) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.586	0.477	0.356
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.805	0.646
	2	0.637	0.524
	3	0.561	0.459
	4	0.429	0.333
	5	0.498	0.381
	6	0.790	0.615

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.948	0.016	0.036	0.927
	2/3	0.892	0.052	0.056	0.848
	3/4	0.874	0.072	0.055	0.827
	4/5	0.898	0.057	0.045	0.857
	5/6	0.945	0.038	0.017	0.920

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.11

Accuracy and Consistency of Classification Indices: Cphn (Grade 11) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.587	0.479	0.360
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.807	0.657
	2	0.643	0.533
	3	0.549	0.444
	4	0.354	0.269
	5	0.541	0.422
	6	0.777	0.602

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.943	0.018	0.038	0.920
	2/3	0.889	0.053	0.058	0.844
	3/4	0.881	0.064	0.055	0.834
	4/5	0.896	0.062	0.042	0.855
	5/6	0.943	0.039	0.019	0.918

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.3.12

Accuracy and Consistency of Classification Indices: Cphn (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.621	0.509	0.374
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.769	0.605
	2	0.697	0.598
	3	0.588	0.486
	4	0.448	0.341
	5	0.559	0.425
	6	0.793	0.587

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.942	0.020	0.037	0.917
	2/3	0.880	0.058	0.062	0.833
	3/4	0.890	0.063	0.047	0.847
	4/5	0.924	0.046	0.030	0.891
	5/6	0.968	0.024	0.008	0.955

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.4.4 Overall

Table 3.4.4.0

Accuracy and Consistency of Classification Indices: Over (Grade K) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.852	0.800	0.691
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.953	0.933
	2	0.752	0.653
	3	0.736	0.635
	4	0.725	0.663
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.952	0.025	0.023	0.933
	2/3	0.954	0.021	0.025	0.935
	3/4	0.964	0.019	0.018	0.948
	4/5	0.982	0.017	0.000	0.982
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.1

Accuracy and Consistency of Classification Indices: Over (Grade 1) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.824	0.756	0.629
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.862	0.774
	2	0.788	0.704
	3	0.875	0.834
	4	0.678	0.586
	5	0.851	0.553
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.969	0.013	0.018	0.956
	2/3	0.927	0.036	0.036	0.897
	3/4	0.945	0.028	0.028	0.921
	4/5	0.983	0.017	0.000	0.982
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.2

Accuracy and Consistency of Classification Indices: Over (Grade 2) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.826	0.764	0.648
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.889	0.822
	2	0.794	0.706
	3	0.884	0.836
	4	0.732	0.672
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.981	0.008	0.011	0.973
	2/3	0.951	0.025	0.024	0.930
	3/4	0.922	0.031	0.047	0.890
	4/5	0.972	0.028	0.000	0.971
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.3

Accuracy and Consistency of Classification Indices: Over (Grade 3) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.811	0.747	0.606
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.916	0.862
	2	0.720	0.608
	3	0.892	0.834
	4	0.723	0.670
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.987	0.005	0.009	0.982
	2/3	0.962	0.022	0.016	0.944
	3/4	0.894	0.034	0.071	0.853
	4/5	0.969	0.031	0.000	0.967
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.4

Accuracy and Consistency of Classification Indices: Over (Grade 4) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.790	0.723	0.581
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.861	0.878
	2	0.675	0.560
	3	0.874	0.801
	4	0.755	0.743
	5	-	-
	6	-	-

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.983	0.008	0.008	0.983
	2/3	0.970	0.019	0.011	0.962
	3/4	0.926	0.030	0.044	0.900
	4/5	0.900	0.100	0.000	0.879
	5/6	0.991	0.009	0.000	0.995

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.5

Accuracy and Consistency of Classification Indices: Over (Grade 5) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.760	0.690	0.547
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.920	0.866
	2	0.682	0.564
	3	0.844	0.769
	4	0.726	0.714
	5	0.575	0.416
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.986	0.005	0.009	0.981
	2/3	0.973	0.015	0.013	0.960
	3/4	0.931	0.029	0.040	0.903
	4/5	0.871	0.116	0.014	0.846
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.6

Accuracy and Consistency of Classification Indices: Over (Grade 6) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.831	0.769	0.660
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.909	0.850
	2	0.744	0.642
	3	0.864	0.808
	4	0.813	0.771
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.981	0.007	0.012	0.973
	2/3	0.958	0.022	0.020	0.940
	3/4	0.921	0.035	0.044	0.889
	4/5	0.972	0.028	0.000	0.967
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.7

Accuracy and Consistency of Classification Indices: Over (Grade 7) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.821	0.754	0.647
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.918	0.865
	2	0.728	0.624
	3	0.857	0.795
	4	0.796	0.754
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.978	0.008	0.014	0.969
	2/3	0.957	0.023	0.020	0.938
	3/4	0.920	0.034	0.046	0.888
	4/5	0.967	0.033	0.000	0.959
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.8

Accuracy and Consistency of Classification Indices: Over (Grade 8) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.815	0.743	0.640
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.913	0.858
	2	0.729	0.625
	3	0.836	0.768
	4	0.799	0.758
	5	-	-
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.973	0.010	0.016	0.963
	2/3	0.953	0.024	0.023	0.933
	3/4	0.922	0.035	0.042	0.890
	4/5	0.966	0.034	0.000	0.957
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.9

Accuracy and Consistency of Classification Indices: Over (Grade 9) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.800	0.722	0.625
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.912	0.857
	2	0.734	0.634
	3	0.827	0.759
	4	0.786	0.727
	5	0.639	0.453
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.971	0.011	0.018	0.960
	2/3	0.951	0.025	0.024	0.930
	3/4	0.928	0.036	0.037	0.898
	4/5	0.951	0.034	0.015	0.933
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.10

Accuracy and Consistency of Classification Indices: Over (Grade 10) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.810	0.734		0.636	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.903		0.842	
	2	0.736		0.636	
	3	0.837		0.775	
	4	0.800		0.740	
	5	0.682		0.492	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.971	0.011	0.018	0.960
	2/3	0.950	0.025	0.025	0.929
	3/4	0.929	0.035	0.035	0.900
	4/5	0.959	0.028	0.012	0.944
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.11

Accuracy and Consistency of Classification Indices: Over (Grade 11) S403 Paper

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.807	0.732		0.633	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.901		0.839	
	2	0.742		0.642	
	3	0.838		0.777	
	4	0.785		0.728	
	5	0.661		0.454	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.971	0.011	0.018	0.959
	2/3	0.948	0.025	0.026	0.927
	3/4	0.929	0.035	0.036	0.900
	4/5	0.959	0.031	0.010	0.946
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

Table 3.4.4.12

Accuracy and Consistency of Classification Indices: Over (Grade 12) S403 Paper

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.848	0.785	0.689
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	1	0.893	0.825
	2	0.771	0.682
	3	0.862	0.814
	4	0.867	0.803
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.971	0.011	0.017	0.960
	2/3	0.943	0.029	0.028	0.919
	3/4	0.934	0.034	0.033	0.907
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Note: Student scale scores and proficiency levels from the state of Colorado and Massachusetts are adjusted for mode effect prior to reporting. As these results are computed differently from other states' results, they are excluded from this analysis.

3.5 Conditional Standard Error of Measurement for Composites

Conditional standard errors of measurement (CSEMs) for the four ACCESS composites provide test users a benchmark of how free the composite scale score is from measurement errors at the various point of the composites. Due to the differential weights applied to different ACCESS domains (see the introduction to this chapter for the weighting conventions), CSEM was estimated using a procedure based on item response theory (IRT) (Lord, 1980) and developed by Price et al. (2006). Price et al. (2006) extended the work by Lord (1980) and Kolen et al. (1992) in estimating the CSEM of a composite consisting of subtests. The basic premise of this procedure is that the student-level CSEM for a weighted composite can be estimated empirically using the IRT-based CSEM for each student on the subtests and the weights associated with the subtests. This method was used to estimate the CSEM for ACCESS composites by treating the ACCESS domains as subtests.

A three-step process was used in deriving the CSEM for ACCESS composites. The derivation was conducted by grade and composite to obtain a unique CSEM for each composite score by grade by composite. Since this procedure relies on empirical student data, which are subject to year-to-year fluctuation, all population student data from the previous ACCESS series were used in the derivation to obtain more stable estimates than using only data from a single series.

Step 1. Since ACCESS domains were calibrated separately, measurement errors associated with each of the ACCESS domains, as expressed in the CSEM, are independent of each other. Therefore, the CSEM for a student with composite score *x*, SEM_x, can be estimated using the equation derived by Price et al. (2006):

$$SEM_x = \sqrt{W_1^2 SEM_1^2 + W_2^2 SEM_2^2 + W_3^2 SEM_3^2 + \cdots + W_k^2 SEM_k^2}$$

Where SEM_i^2 is the student's IRT-based score error variance or student's squared CSEM in ACCESS domain *i* and W_i is the weight applied to domain *i*, for *i*=1,…,*k*.

Step 2. Due to the differential weights applied to different ACCESS domains, two students with the same sum of weighted domain score, or composite, may obtain different CSEMs; therefore, an additional step was taken to obtain a unique value for each composite score. Specifically, the expected value of the CSEM functions for a composite score was estimated using a regression approach, and this expected value is used as the reported CSEM for that composite score.

Step 3. A linear smoothing procedure was applied to derive the CSEM for composite scores that were not observed in the data.

The figures in this section show graphically the CSEM for various composite scores by grade level. Figures show the relationship between the students' composite scores on the horizontal axis and conditional measurement errors on the vertical axis. Each point in the figures represents a student in the dataset, expressing both the student's CSEM and that student's scale score for the given composite score. Values for students who received the lowest possible scores on any

ACCESS domains are not plotted, as the conditional measurement errors for these students cannot be computed accurately. For grade-level clusters with multiple grades, different colors are used in the figures to represent students in different grades.

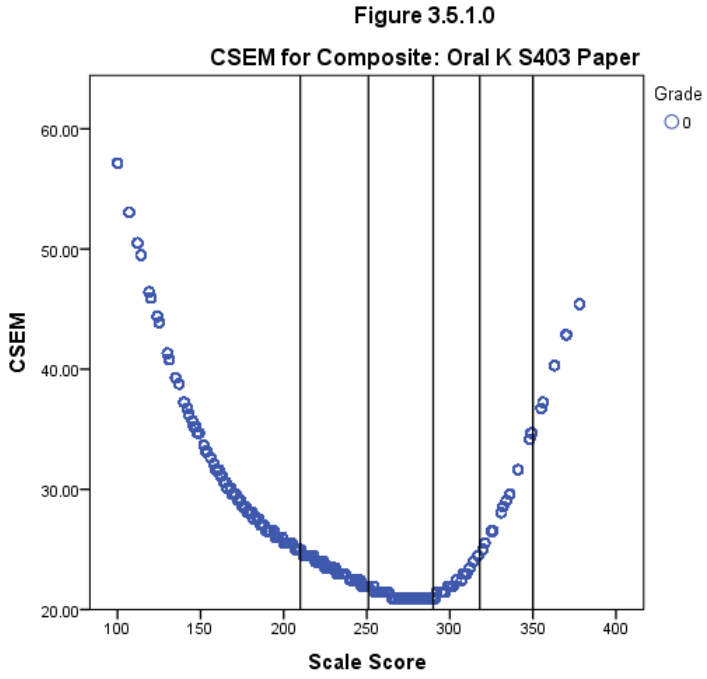
Five vertical lines in the figure indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1–6) for the composites.

Low CSEM values indicate less measurement error or more accuracy in measurement. The general trend in these figures shows that the CSEMs are lower and fairly constant in the middle of the score range and higher and more variable for extreme low and high scores, as expected.

3.5.1 Oral

3.5.1.0 Kindergarten

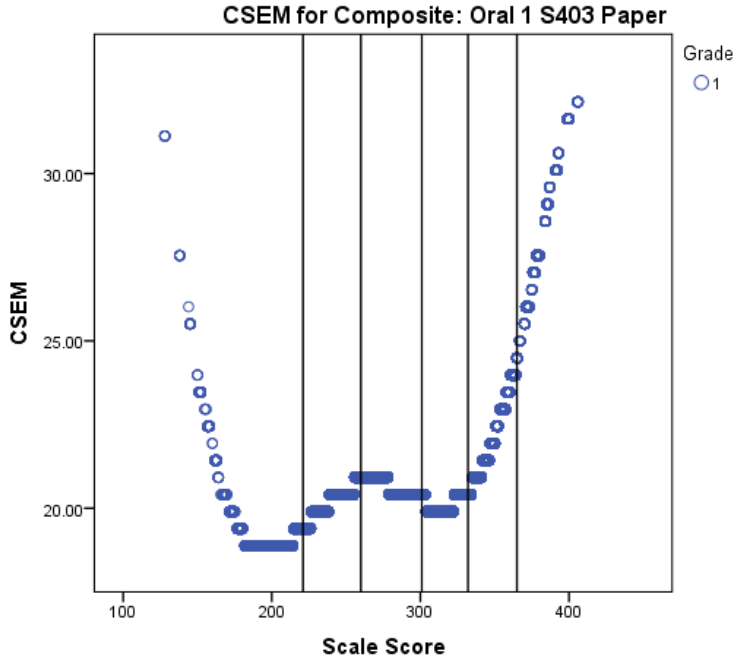
Figure 3.5.1.0



Note: Grade 0 is Kindergarten

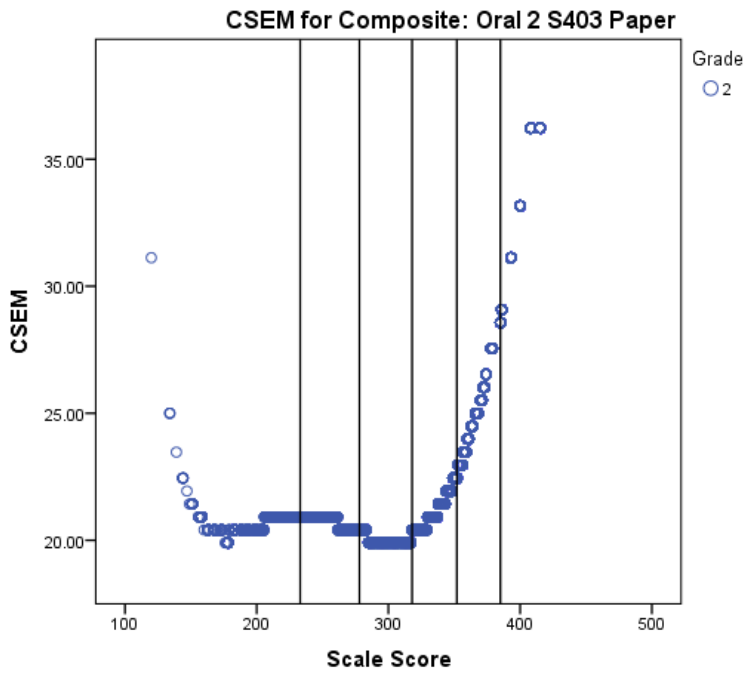
3.5.1.1 Grade 1

Figure 3.5.1.1



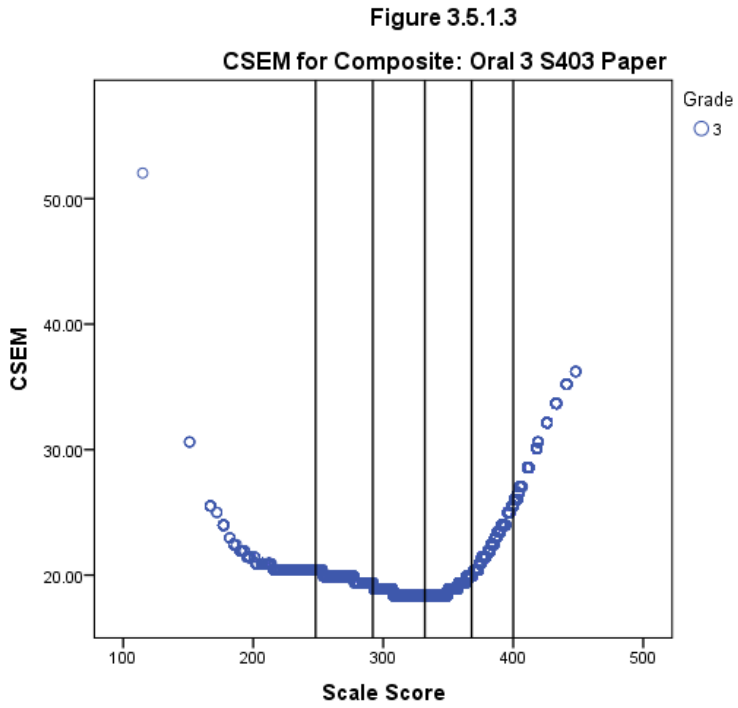
3.5.1.2 Grade 2

Figure 3.5.1.2



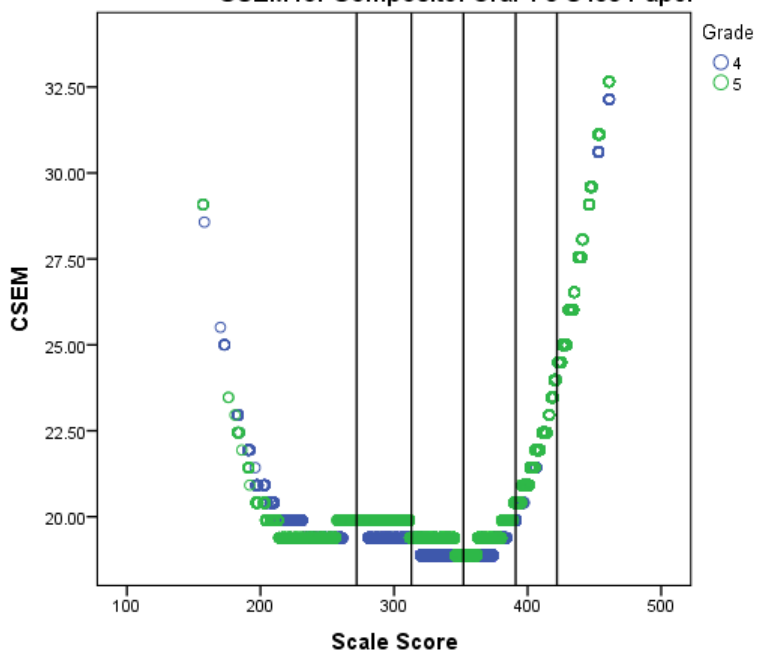
3.5.1.3 Grade 3

Figure 3.5.1.3



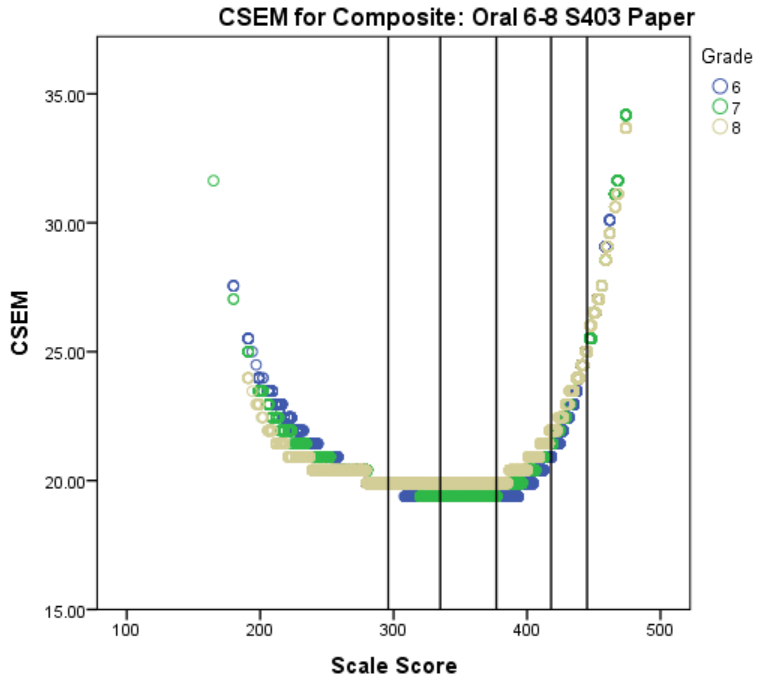
3.5.1.4 Grades 4–5

Figure 3.5.1.4



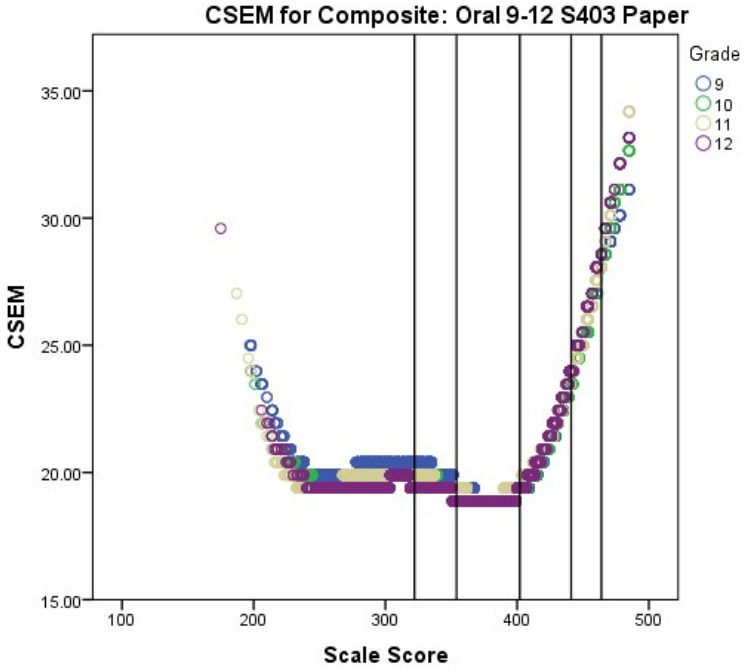
3.5.1.5 Grades 6–8

Figure 3.5.1.5



3.5.1.6 Grades 9–12

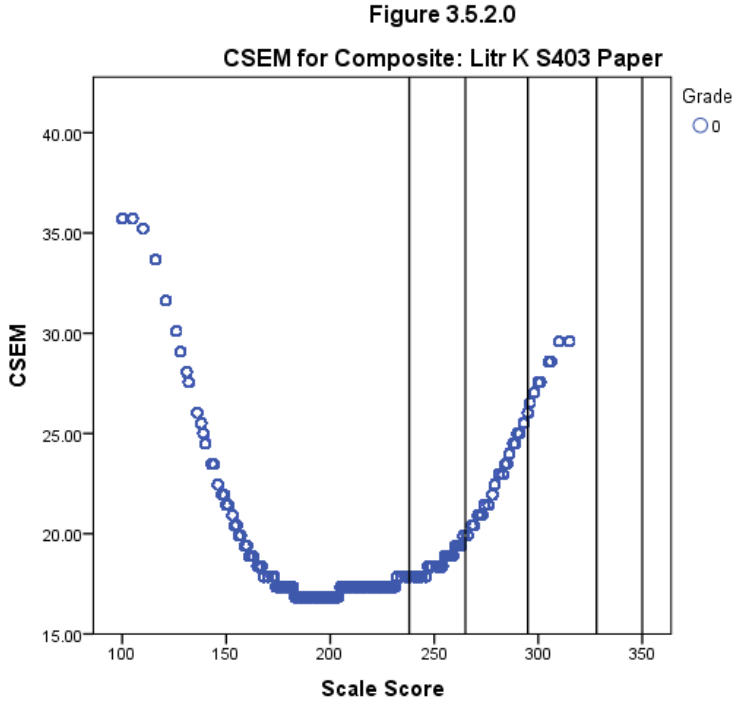
Figure 3.5.1.6



3.5.2 Literacy

3.5.2.0 Kindergarten

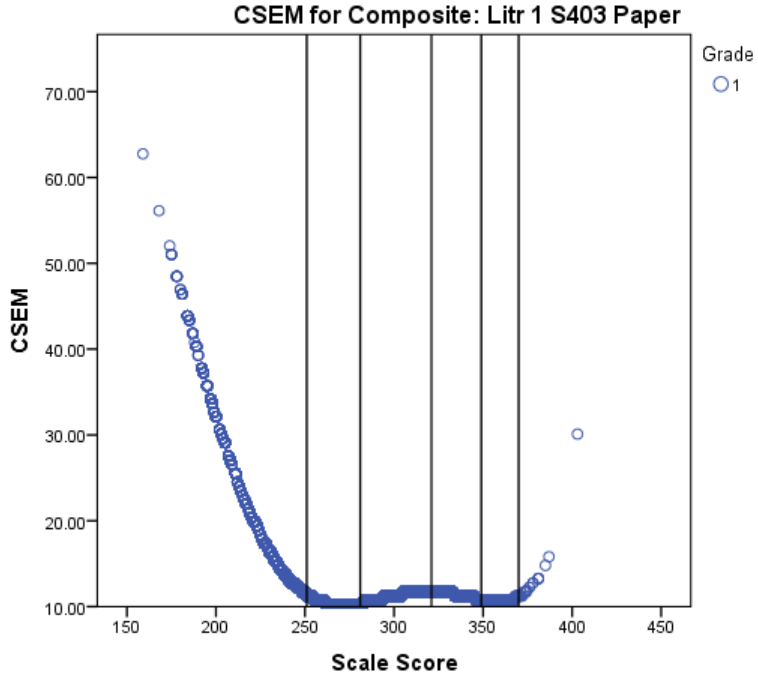
Figure 3.5.2.0



Note: Grade 0 is Kindergarten

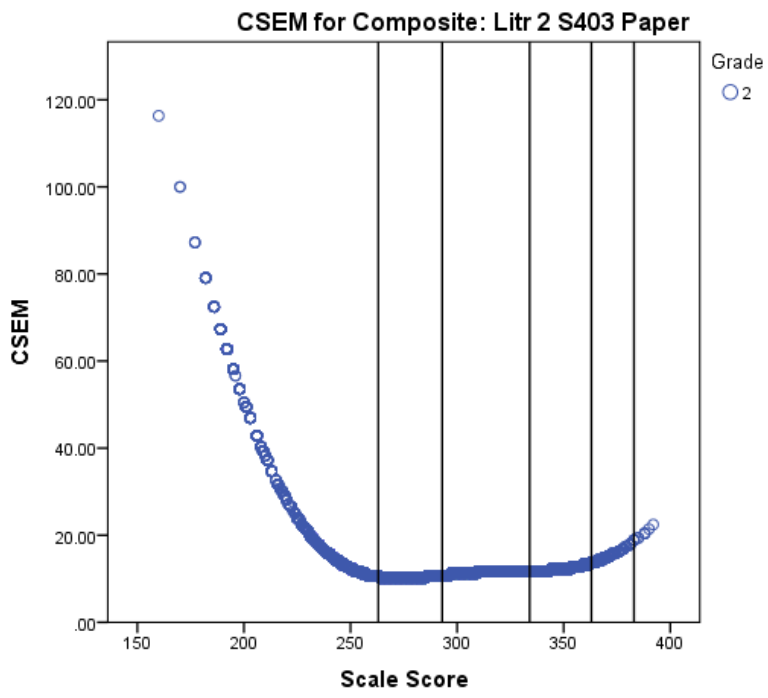
3.5.2.1 Grade 1

Figure 3.5.2.1



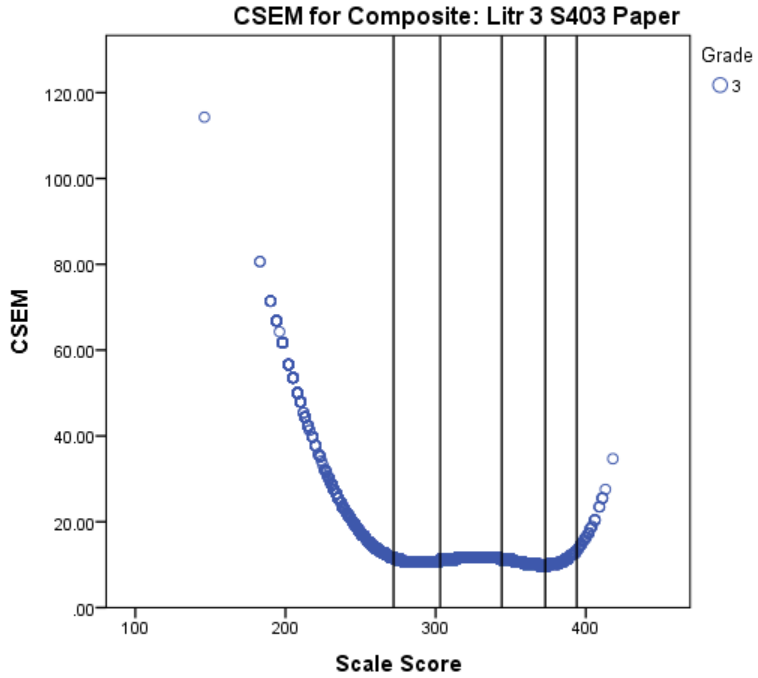
3.5.2.2 Grade 2

Figure 3.5.2.2



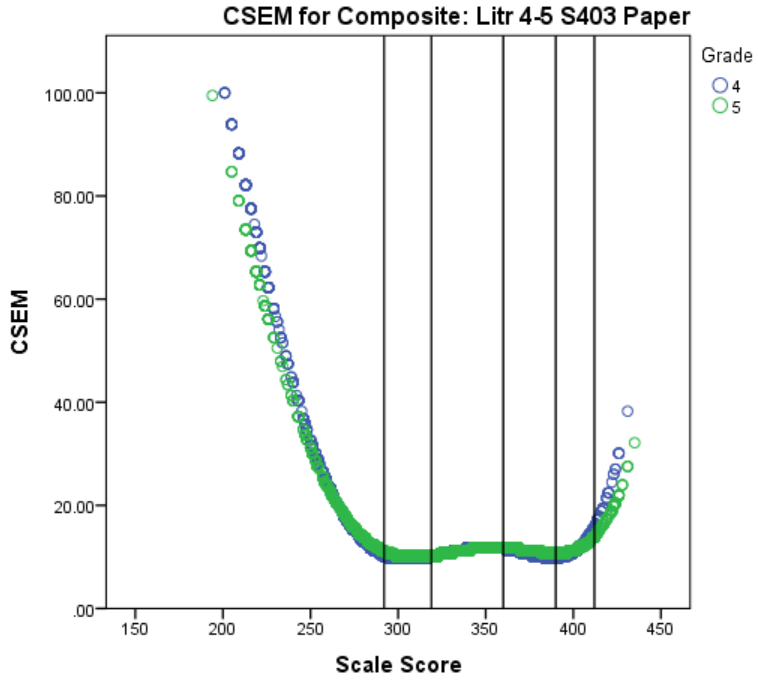
3.5.2.3 Grade 3

Figure 3.5.2.3



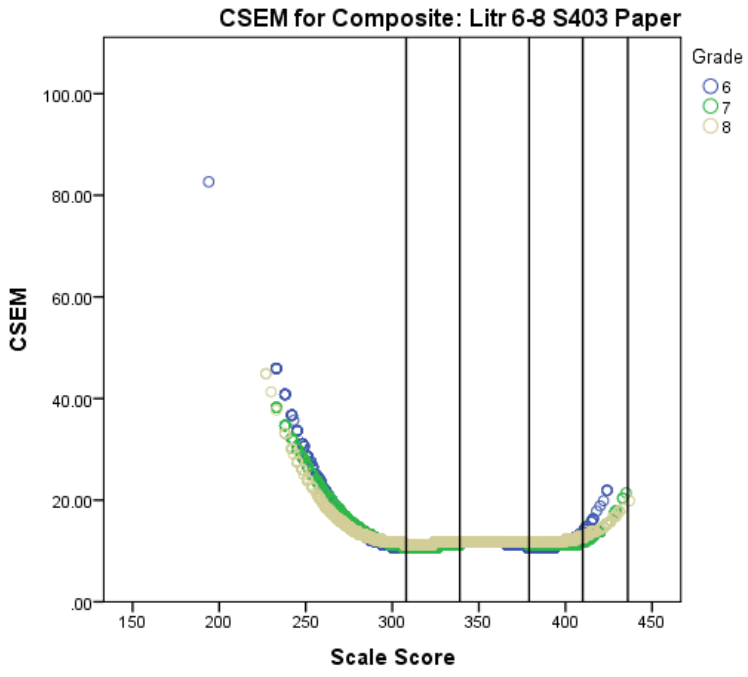
3.5.2.4 Grades 4–5

Figure 3.5.2.4



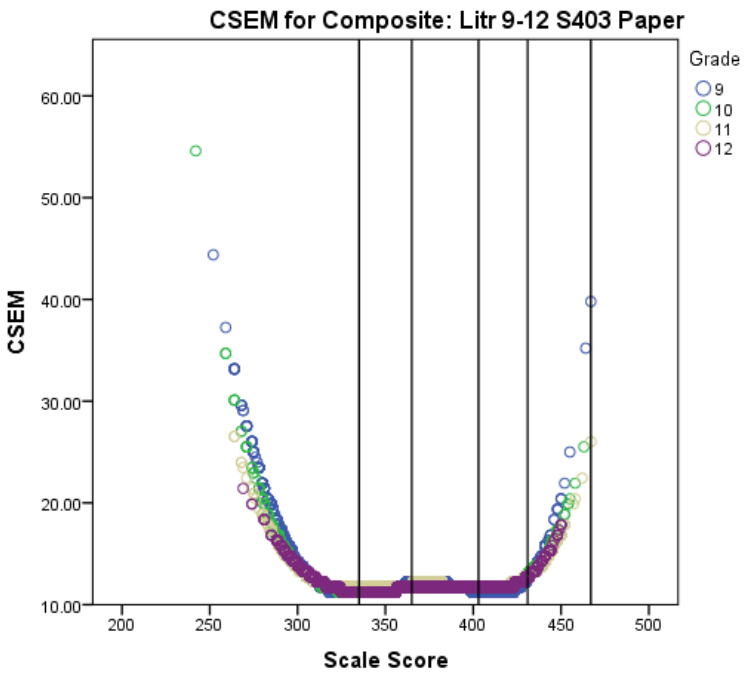
3.5.2.5 Grades 6–8

Figure 3.5.2.5



3.5.2.6 Grades 9–12

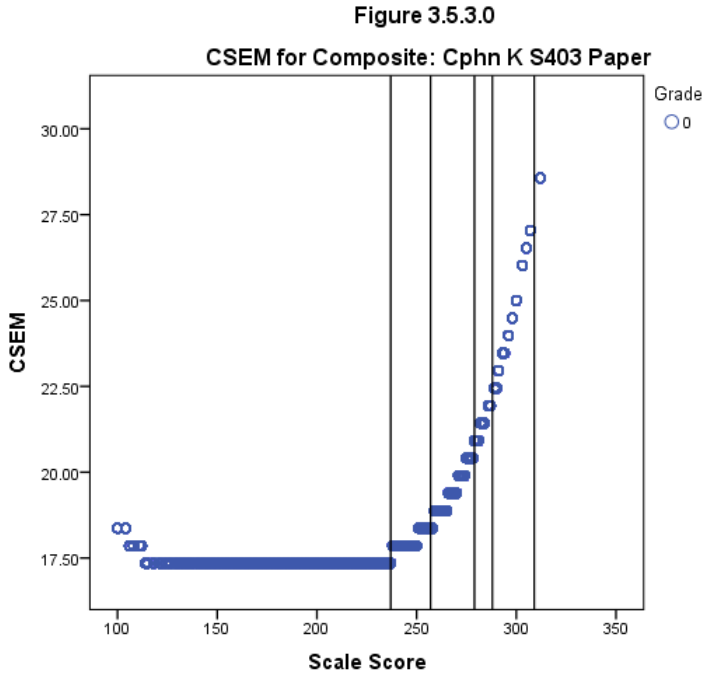
Figure 3.5.2.6



3.5.3 Comprehension

3.5.3.0 Kindergarten

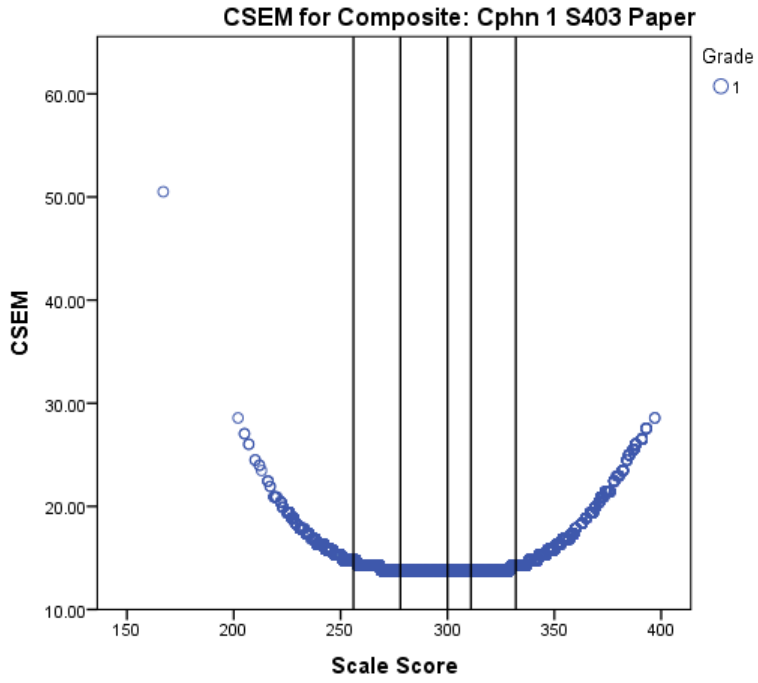
Figure 3.5.3.0



Note: Grade 0 is Kindergarten

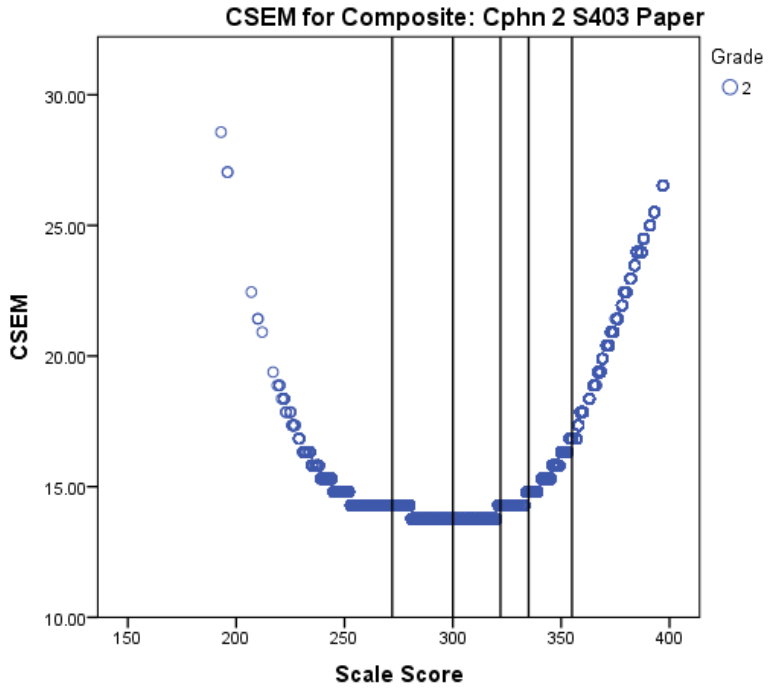
3.5.3.1 Grade 1

Figure 3.5.3.1



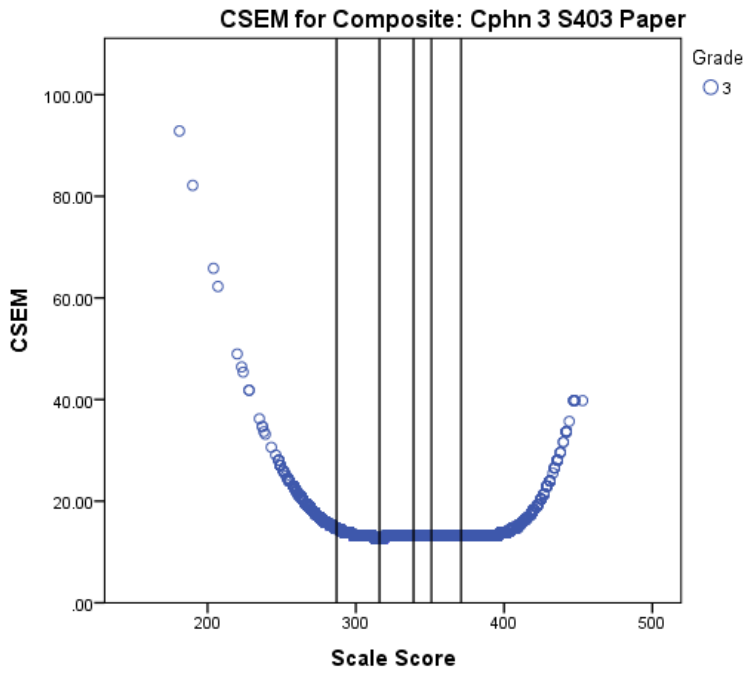
3.5.3.2 Grade 2

Figure 3.5.3.2



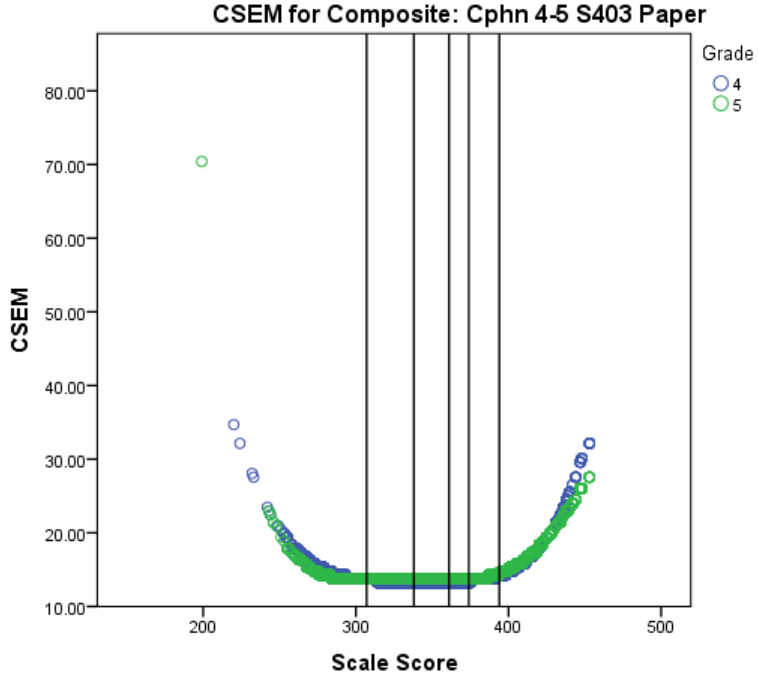
3.5.3.3 Grade 3

Figure 3.5.3.3



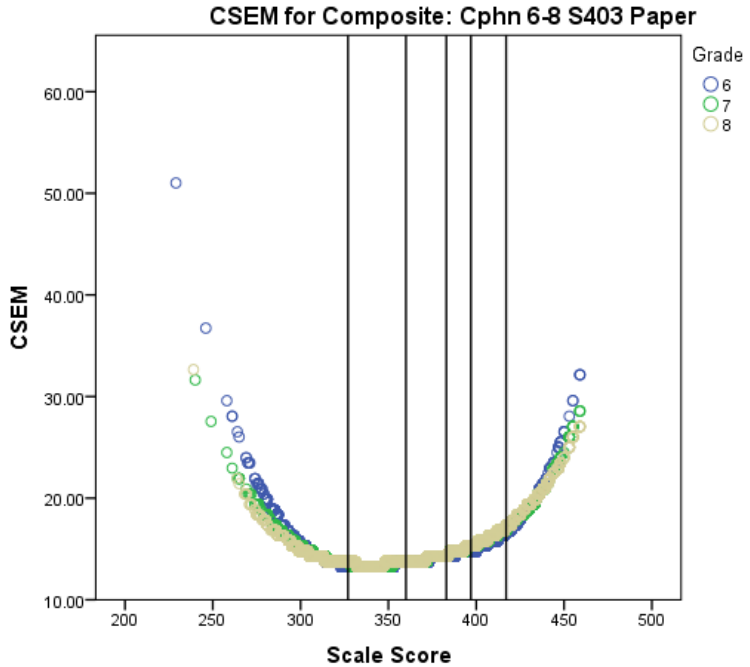
3.5.3.4 Grades 4–5

Figure 3.5.3.4



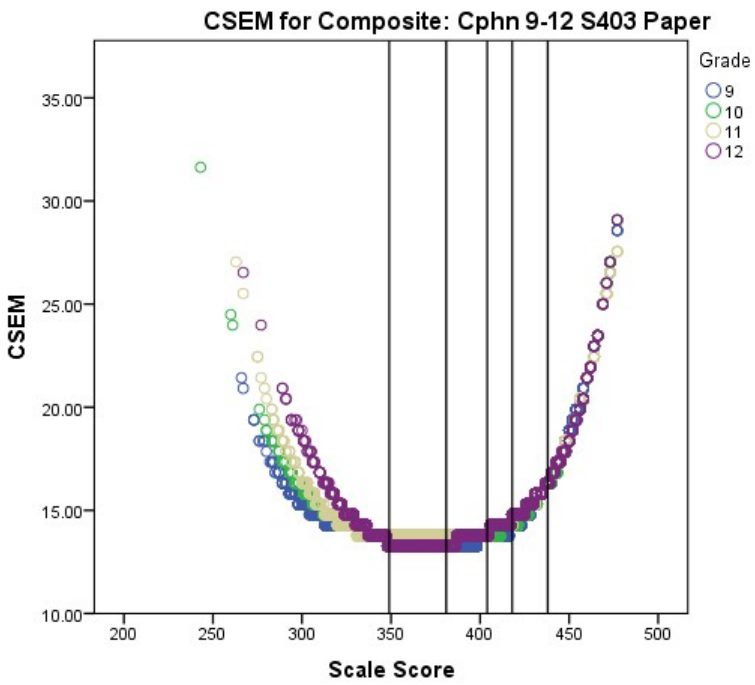
3.5.3.5 Grades 6–8

Figure 3.5.3.5



3.5.3.6 Grades 9–12

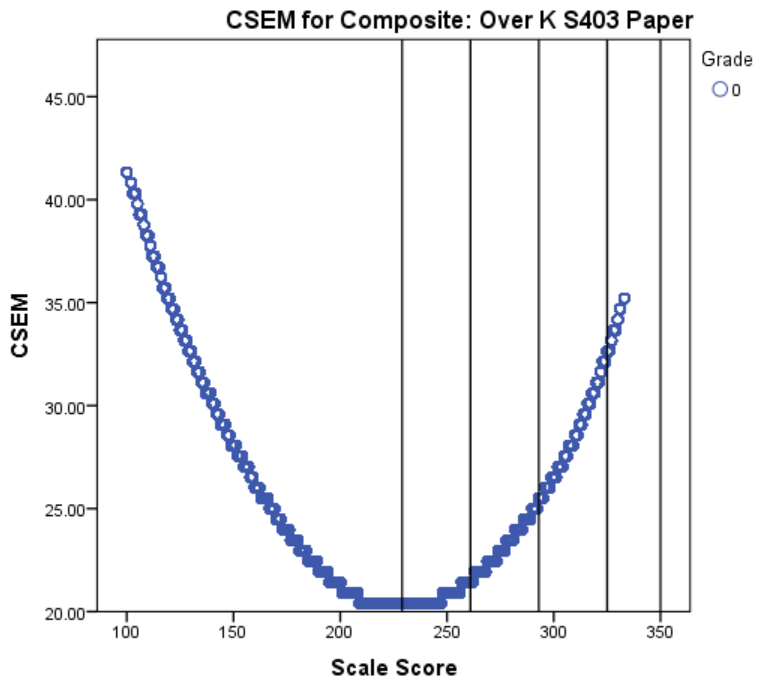
Figure 3.5.3.6



3.5.4 Overall

3.5.4.0 Kindergarten

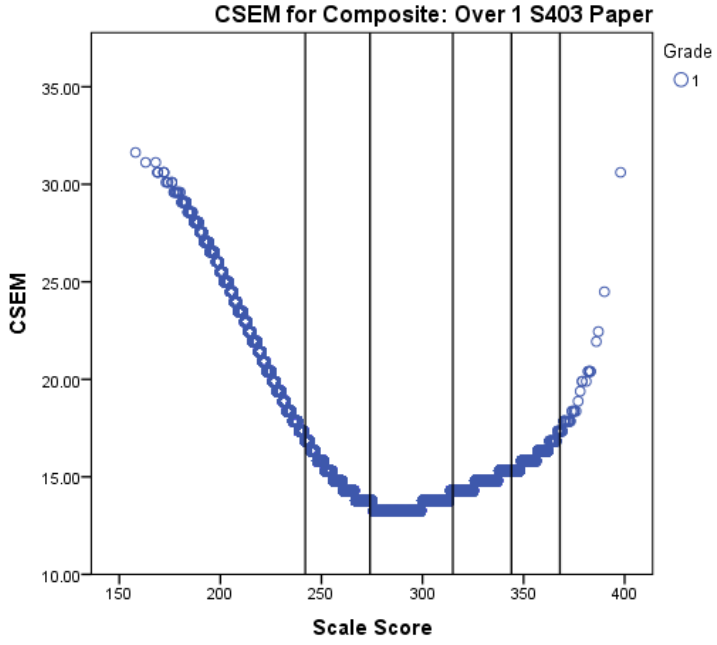
Figure 3.5.4.0



Note: Grade 0 is Kindergarten

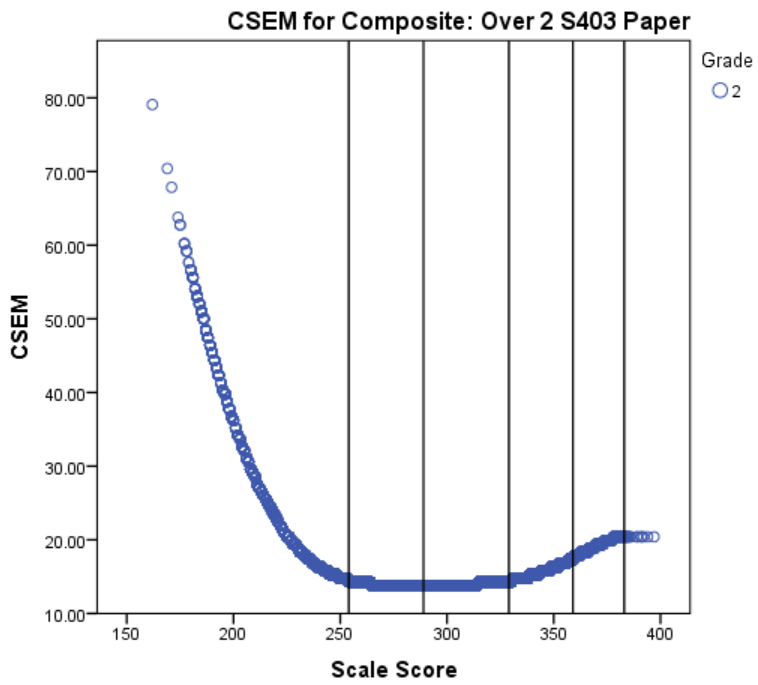
3.5.4.1 Grade 1

Figure 3.5.4.1



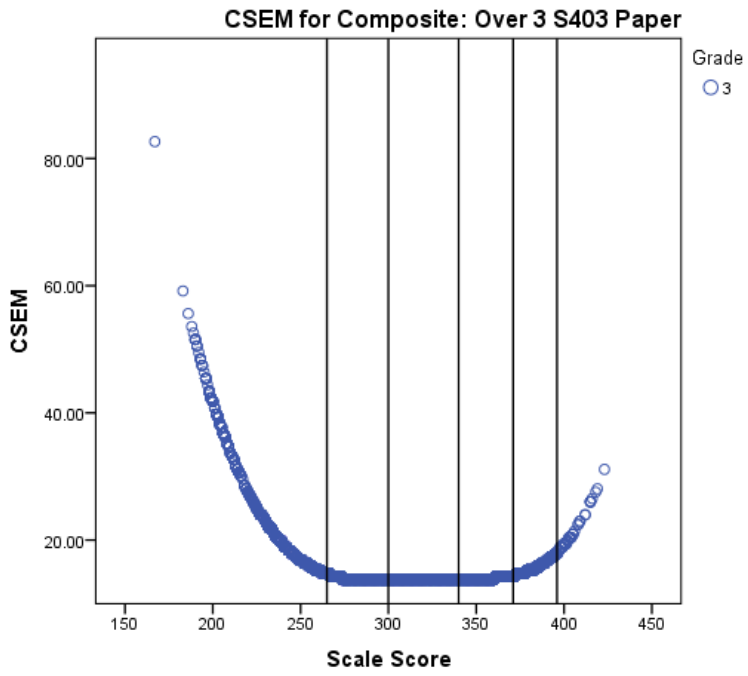
3.5.4.2 Grade 2

Figure 3.5.4.2



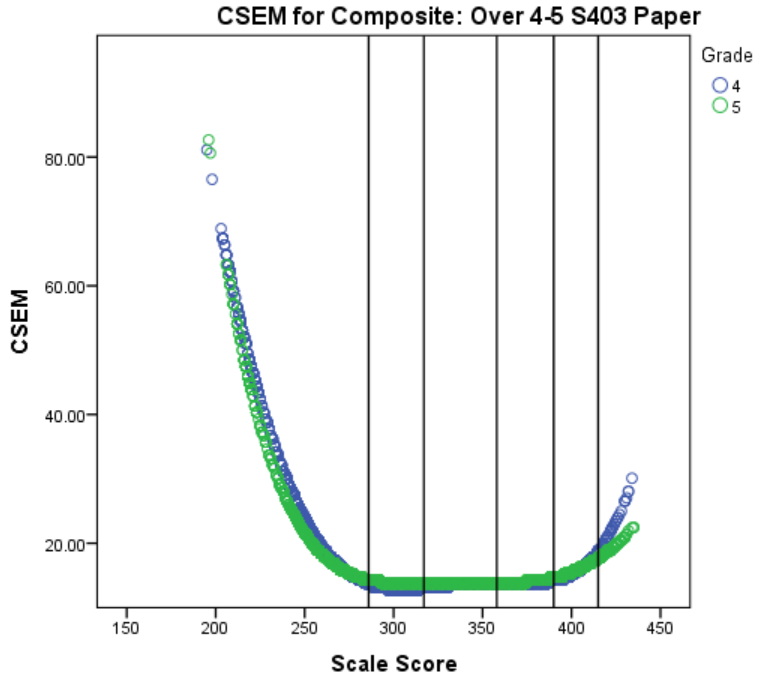
3.5.4.3 Grade 3

Figure 3.5.4.3



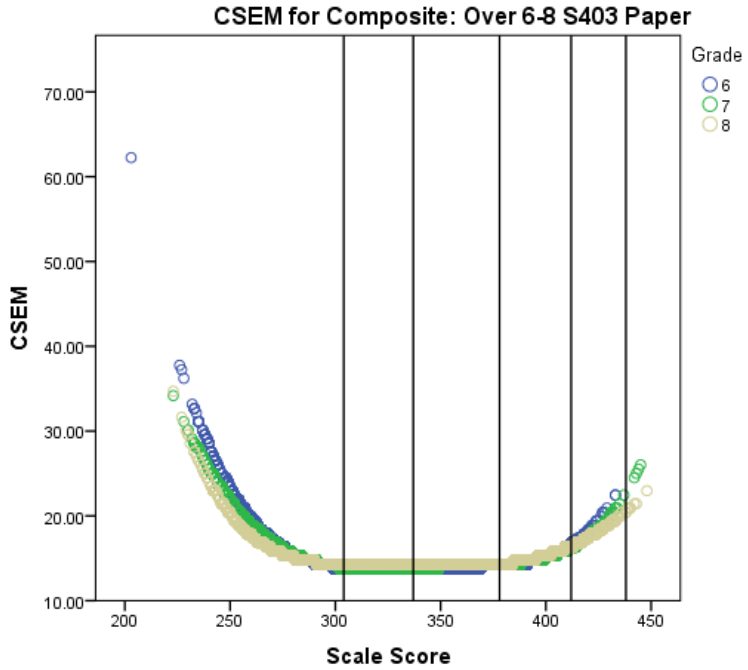
3.5.4.4 Grades 4–5

Figure 3.5.4.4



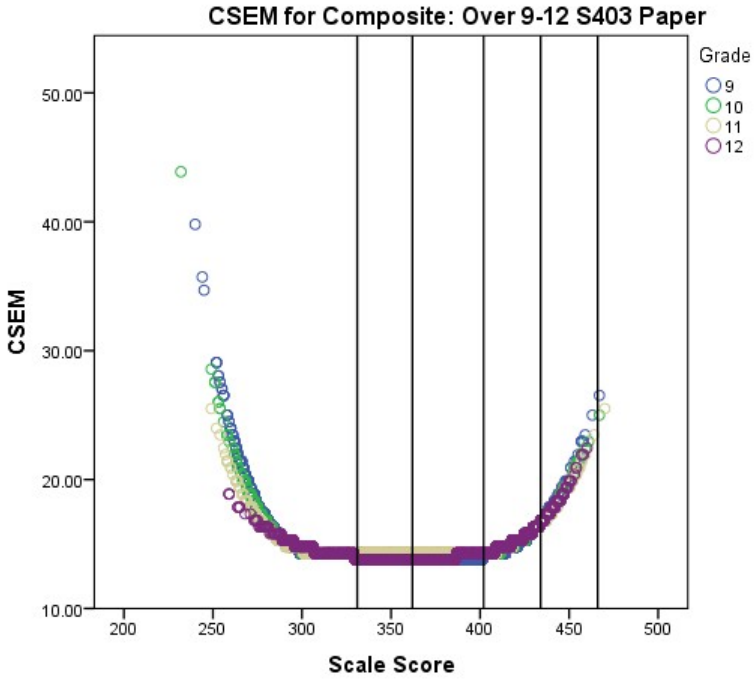
3.5.4.5 Grades 6–8

Figure 3.5.4.5



3.5.4.6 Grades 9–12

Figure 3.5.4.6



4. Annual Updates of Validity Evidence

According to the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014), validity is the degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed use.

Particular interpretations for specified uses begin by specifying the construct the test is intended to measure. Validity is the degree to which all the accumulated evidence supports the intended interpretation of the test scores for the proposed use. Rather than referring to distinct types of validity, the *Standards* refer to types of validity evidence.

According to the *Standards* the evidence can be based on 1) test content, 2) response processes, 3) internal structure, and 4) relations to other variables.

4.1. Standards

4.1.1. Test content

Important validity evidence can be obtained from an analysis of the relationship between the content of a test and the construct it is intended to measure. Test content refers to the themes, wording, and format of the items, tasks, or questions on a test. Administration and scoring may also be relevant to content-based evidence. Evidence based on test content can include logical or empirical analyses of the adequacy with which the test content represents the content domain and of the relevance of the content domain to the proposed interpretation of test scores. Evidence based on test content can also come from expert judgement of the relationship between parts of the test and content.

4.1.2. Response processes

Theoretical and empirical analyses of the response processes of test takers can provide evidence concerning the fit between the construct and the detailed nature of the performance or response actually engaged in by test takers. Evidence based on response processes generally comes from analysis of individual responses. Evidence of response processes can contribute to answering questions about differences in meaning or interpretation of test scores across relevant sub-groups of test takers. Studies of response processes are not limited to the test taker. Assessment often relies on observers or judges to record and/or evaluate test takers' performances or products.

4.1.3. Internal structure

Analyses of the internal structure of a test can indicate the degree to which the relationships among the test items and test components conform to the construct on which the proposed test score interpretations are based. The conceptual framework for a test may imply a single

dimension of behavior, or it may posit several components that are each expected to be homogeneous.

4.1.4. Relations to other structure

In many cases, the intended interpretation for a given use implies that the construct should be related to some other variables, and as a result, analysis of the relationship of the scores to variables external to the test provides another important source of validity evidence. Evidence about relations to other variables is also used to investigate questions of differential prediction for subgroups. In the test-criterion relationship the fundamental question is, how accurately do test scores predict criterion performance? Historically, two designs, often called predictive and concurrent, have been differentiated for evaluating test-criterion relationships. A predictive study indicates the strength of the relationship between test scores and criterion scores that are obtained at a later time. A concurrent study obtains test scores and criterion information at about the same time.

4.2. Annual validity studies

4.2.1. ACCESS for ELLs 2.0 Series 400 Construct Validity Study

Grant, R. & Kraninger, N. (2019). *ACCESS for ELLs 2.0 Construct Validation Study - ACCESS for ELLs 2.0, Series 400 with 2015-2016 Data*. WIDA Technical Report TR-2019-1 (unpublished).

A research study was conducted to examine the construct validity of the internal structures of the paper and online versions of ACCESS for the grade-level cluster forms within Grades 1-12 through construct validity techniques at the item level. The research questions were:

1. Are there 4 distinct, orthogonally related, first-order academic English language factors (Listening, Reading, Speaking, and Writing) underpinning English Learners' responses to Tiers A, B and C items within the grade-level cluster forms of the paper and online versions of the assessment?
2. If so, is there a 2nd-order, underlying factor (academic English language) orthogonally related to each first-order factor within each tier?

Confirmatory factor analyses were used to investigate the first research question and structural equation modeling techniques were used to investigate the second question. All 33 models had satisfactory model fit statistics. The conclusion was that the ACCESS for ELLs 2.0, Series 400 assessment (online and paper) has an internal structure consistent with four academic English language domains (Listening, Reading, Writing and Speaking) and a second order underlying language factor (academic English language). This suggests that it is appropriate to report separate English language domain scores and an overall composite score for the paper and online versions of the test.

4.2.2. WIDA Screener Online and WIDA ACCESS for ELLs: Examining the Relationship between Student Scores on Two Assessments

MacGregor, D., and Sahakyan, N. (in preparation). *WIDA Screener Online and WIDA ACCESS for ELLs: Examining the Relationship Between Student Scores on Two Assessments*. WIDA Technical Brief.

WIDA conducted a comparability study between WIDA Online Screener (hereafter Screener) and WIDA ACCESS (hereafter ACCESS). The purpose of the study was to understand to what extent scores on Screener predict scores on ACCESS, and how that relationship is potentially affected by several covariates, such as grade level, time between tests, and IEP status.

The results of this analysis show that Screener scores are highly correlated with and strongly predictive of ACCESS scores controlling for individual-level and institutional factors. Controlling for all of the above-mentioned factors, a one-point increase in the Screener Overall Composite Scale Score is associated with about a 0.70 point increase in the ACCESS Overall Composite Scale. Students with special needs score about 8 scale score points lower on ACCESS, compared to non-IEP students. Hispanic ELs score about 5 points below non-Hispanics, while female ELs score about 2.3 scale score points higher compared to their male counterparts. Finally, parameter estimates indicate that in addition to individual-level covariates, there are statistically significant grade, school, district and state effects; however, the predictive relationship between the two tests is not altered when these factors are accounted for.

These results provide evidence for the predictive validity of Screener for schools in the WIDA Consortium. The stable and strong predictive relationship between Screener and ACCESS Overall composite scores suggests that schools can confidently use the WIDA Screener Overall composite score when making placement decisions for ELLs. In addition, they provide validity evidence for the use of Screener as an identification tool for students in the WIDA Consortium.

4.2.3. 2018 ACCESS for ELLs 2.0 Speaking Study

Bishop, K., Read, S., Gocer-Sahin, S., and Akanda, M. (2019). *ACCESS for ELLs Speaking study*. WIDA Technical Report.

Although the amount of score difference varied year to year, in the domain of Speaking, paper scores have been consistently higher than online counterpart. The purpose of this study was to investigate potential differences in the test-taking and scoring procedures between the paper Speaking test and the online Speaking test. The results of this study are related to content validity since it explores possible reasons of differences among different administration and scoring procedures (AERA, APA & NCME, 2014) and evidence based on response processes.

In this study a WIDA researcher used the Speaking scores of ACCESS 401 population data. Data was collected from IL and FL. The WIDA researcher went to each school site to set up the

speaking test sessions and met teachers to administer and score the test. The student's entire testing session and spoken responses were audio recorded by the WIDA researcher.

A quantitative and qualitative analysis was conducted. As the first step of the quantitative analysis, descriptive statistics such as raw score distributions, means across states, tiers, and raters were calculated. Second, to explore the rater effects of test administrative modes, Patz's (1996) hierarchical rater model (HRM) for polytomous Speaking rating data scored by multiple raters to scale examinees and items was applied to model aspects of consensus among raters, and to model individual rater severity and consistency effects (Patz, Junker, Johnson, and Mariano, 2002).

Qualitative analysis was conducted in four steps.

1) Interviews: After the WIDA researcher recorded students' Paper speaking test responses, we interviewed paper raters for their reasoning of scores. For online raters, the interview questions were sent to them and after their scorings were done, they typed their answers to the survey questionnaire document.
2) Categorizing interviews and observations: Based on interviews and the observations from administering tests and testing environment, the researcher categorized the interview comments into themes.
3) FL score comparison: one WIDA researcher, a rater trainer in FL scored students' responses as she was observing the paper testing. The WIDA researcher's scores and FL local raters' scores and online raters' score in FL were compared.
4) Score difference examination: two WIDA researchers examined paper and online ratings where there were score differences for 50 tasks to explore possible reasons of score differences.

The quantitative analysis results, HRM results, showed us that online raters are more consistent in assigning the same score to work of the same quality than paper raters. Tier B/C is more difficult than Tier A. Interestingly, although number of ideal scoring is three for paper raters for tier A, it is six for tier B/C. Similar to paper raters, in tier A, number of ideal scoring of online raters is seven, however it is eight for tier B/C. Although there isn't much change for online raters, in general it can be said that when test becomes more difficult, raters become more reliable in scoring.

Based on the local observations and rater reliability results, the WIDA Researcher has identified a need for additional clarification of the various aspects of the Speaking domain that should be provided in local and online trainings. It can be said that there is no one reason to give a high or low score. However, WIDA Researcher observed that, word choice was overlooked as sophistication. Even if sophistication was not present in the answer, online raters scored 3. Even though one of three criteria was missing, both raters still gave score 3 no matter which criterion is missing.

It was also observed that sometimes teacher prompting occurred, but they were within legitimate promoting options according to test administrator manual. Occasionally some prompting was very guided and specific, but student answers did not get more sophisticated. Teacher-prompted answers did not always lead to higher scores by paper raters. Some paper raters paid more attention to criteria outside of the rubric such as “complete sentences” and scored lowered. Interpretation of Nina model to the rubric description is an issue. For example, for P3 or P5 questions, Nina model answer has detailed explanations. It is not clear whether this explanation is referring to sophistication level in the rubric.

References

Allen, N. L., Carlson, J. E., & Zalanak, C. A. (1999). *The NAEP 1996 technical report.* Washington, DC: National Center for Education Statistics.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Brennan, R. (2004). Linking with equivalent group or single group design (LEGS) (Version 2.0) [Computer software]. Iowa City, IA: Center for Advanced Studies in Measurement and Assessment.

Center for Applied Linguistics. (2016). *ACCESS for ELLs® Series 400 Listening and Reading scale maintenance: Technical brief.* Madison, WI: WIDA Consortium.

Center for Applied Linguistics. (2017). *ACCESS for ELLs® 2.0 Speaking and Writing score scale reconstruction: Technical brief.* Washington, DC: Author.

Center for Applied Linguistics. (2018). *Annual technical report for ACCESS for ELLs® English Language Proficiency Test, Series 401 Paper, 2016–2017 administration* (WIDA Consortium Annual Technical Report No. 13B). Washington, DC: Author.

Cook, H. G., & MacGregor, D. (2017). *The ACCESS for ELLs 2.0 2016 Standard setting study* (Technical Report). Madison, WI: Board of Regents of the University of Wisconsin System.

Elementary and Secondary Education Act of 1965, amended 2015. 20 USC §6301-8961.

Gottlieb, M. (2004). *English Language Proficiency Standards for English language learners in Kindergarten through Grade 12: Framework for large-scale state and classroom assessment.* Madison, WI: WIDA Consortium.

Kamata, A., Turhan, A., & Darandari, E. (2003, April). *Estimating reliability for multidimensional composite scale scores.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Kane, M., & Case, S. M. (2004). The reliability and validity of weighted composite scores. *Applied Measurement in Education, 17,* 221–240.

Kenyon, D. M. (2006). *Development and field test of ACCESS for ELLs*® (WIDA Consortium Technical Report No. 1). Washington, DC: Center for Applied Linguistics.

Kenyon, D. M., Ryu, J. R., & MacGregor, D. (2013). *Setting grade level cut scores for ACCESS for ELLs*® (WIDA Consortium Technical Report No. 4). Washington, DC: Center for Applied Linguistics.

Kolen, M. J., Hanson, B. A., & Brennan, R. L. (1992). Conditional standard errors of measurement. *Journal of Educational Measurement, 29,* 285–307.

Lee, W., Hanson, B. A., & Brennan, R. L. (2002). Estimating consistency and accuracy indices for multiple classifications. *Applied Psychological Measurement, 26,* 412–432.

Linacre, J. M. (2002, Autumn). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*(2), 878. Retrieved from http://www.rasch.org/rmt/rmt162f.htm

Linacre, J. M. (2006). Winsteps Rasch analysis (Version 3.60.1) [Computer software]. Retrieved from http://www.winsteps.com

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32,* 179–197.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NO: Lawrence Erlbaum Associates.

MacGregor, D., Kenyon, D. M., Gibson, S., & Evans, E. (2009). *Development and field test of Kindergarten ACCESS for ELLs®*. Madison, WI: WIDA Consortium.

Mantel, N., & Haenszel, W. (1959). Statistical aspect of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719–748.

Meyer, J. P. (2018). jMetrik [Computer software]. Retrieved from http://itemanalysis.com/jmetrik-download/

Patz, R. J. (1996). *Markov chain Monte Carlo methods for item response theory models with applications for the National Assessment of Educational Progress*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Patz, R. J., Junker, B. W., Johnson, M. S., & Mariano, L. T. (2002). The hierarchical rater model for rated test items and its application to large-scale educational assessment data. *Journal of Educational and Behavioral Statistics, 27*(4), 341–384. doi:10.3102/10769986027004341

Price, L. R., Lurie, A., Raju, N., Wilkins, C., & Zhu, J. (2006). Conditional standard errors of measurement for composite scores on the Wechsler Preschool and Primary Scale of Intelligence – Third Edition. *Psychological Reports, 98*(1), 237–252.

Rudner, L. (2001, Spring). Informed test component weighting. *Educational Measurement: Issues and Practice, 20*(1), 16–19.

Waller, N. G. (n.d.). EZDIF: A computer program for detecting uniform and nonuniform differential item functioning with the Mantel-Haenszel and logistic regression procedures [Computer software]. Davis, CA: University of California – Davis.

WIDA Consortium. (2007). *English language proficiency standards and resource guide, 2007 edition, PreKindergarten through Grade 12*. Madison, WI: Board of Regents of the University of Wisconsin System.

WIDA Consortium. (2012). *2012 amplification of the English language development standards Kindergarten–Grade 12*. Madison, WI: Board of Regents of the University of Wisconsin System.

Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement.* Chicago, IL: MESA Press.

Young, M. J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment* (CSE Technical Report 475). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education and Information Studies.

Zieky, M. (1993). DIF statistics in test development. In P.W. Holland and H. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale, NJ: Erlbaum.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30,* 233–251.

Acknowledgments

We would like to extend our appreciation to the many CAL and WIDA staff members who have supported this work.

From CAL:

Keira Ballantyne, Ph.D.
Tanya Bitterman, M.A.
Yage (Leah) Guo, Ph.D.
Bernadette Jerome, M.S.
Michele Kawood, M.S.Ed.
Justin Kelly, Ph.D.
Dorry M. Kenyon, Ph.D.
Tristan Kirkman, M.A.
Mohammed Louguit, Ph.D.
Erin Shaw-Meadow, M.Sc.
Samantha Musser, M.A.
Yu-Chia Wu, M.Ed.
Shu Jing Yen, Ph.D.
Xin Yu, M.A.

From WIDA:

Mohammad Akanda, Ph.D.
Kyoungwon Bishop, Ph.D.
Sakine Göçer Sahin, Ph.D.